

OPERATIONAL MODEL FOR EMPTY CONTAINER REPOSITIONING

LONG YIN

NATIONAL UNIVERSITY OF SINGAPORE

2012

**OPERATIONAL MODEL FOR EMPTY CONTAINER
REPOSITIONING**

LONG YIN

(B.Eng., Shanghai Jiao Tong University)

**A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF INDUSTRIAL AND SYSTEMS
ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE**

2012

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Long Yin, 20 Aug 2012

LONG YIN

20 AUG 2012

ACKNOWLEDGEMENTS

I would like to thank all the people who have helped and inspired me during my doctoral study.

First and foremost, I would like to express my deepest appreciation to my supervisors: A/Prof. Lee Loo Hay and A/Prof. Chew Ek Peng, for their valuable guidance during my research and study. Their perpetual energy and enthusiasm in research had motivated me, even during tough times in my PhD pursuit. Their contributions of time and ideas make research life stimulating and rewarding for me.

The members of maritime logistics and supply chain systems research group have also contributed immensely to me. I am grateful for the project collaborators on empty container repositioning, Luo Yi and Shao Jijun, for their friendships as well as good advices and collaboration throughout the project and my research life.

I also wish to thank the scholarship support from department of Industrial and Systems Engineering in National University of Singapore, without which this thesis would never have been written. Gratitude also goes to all other faculty members and staffs in the department of Industrial and Systems Engineering, especially the members of Systems and Modeling and Analysis Lab, for their supports and advices.

Finally, I would like to thank my families, especially my husband Yao Zhishuang, for their continuous support, confidence and constant love on me.

Long Yin, 20 Aug 2012

LONG YIN

TABLE OF CONTENTS

DECLARATION	i
ACKNOWLEDGEMENTS.....	ii
TABLE OF CONTENTS	iii
SUMMARY	viii
LIST OF TABLES	x
LIST OF FIGURES.....	xi
LIST OF SYMBOLS	xii
Chapter 1 INTRODUCTION.....	1
1.1 Background.....	3
1.1.1 Overview of empty container repositioning operation in shipping industry	3
1.1.2 Uncertainties in maritime empty container repositioning	5
1.2 Research scope and objectives	6
1.3 Organization of thesis	8
Chapter 2 LITERATURE REVIEW.....	10
2.1 Empty container repositioning problem	10
2.1.1 Strategic level empty container repositioning	10
2.1.2 Tactical level empty container repositioning.....	12
2.1.3 Operational level empty container repositioning	16
2.1.4 Empty container repositioning with uncertainty	21
2.2 Methods to solve stochastic empty container repositioning	

problem	24
2.2.1 General methods for stochastic fleet management problem	24
2.2.2 Sample average approximation method	28
2.2.3 Scenario decomposition for the stochastic problem with multiple scenarios	32
2.2.4 Sampling schemes to enhance the performance of sample average approximation	34
Chapter 3 A TIME SPACE NETWORK MODEL ON EMPTY CONTAINER FLOW MANAGEMENT	38
3.1 Problem description	38
3.1.1 General decision process of empty container repositioning	39
3.1.2 Time space network	41
3.2 Mathematical model	43
3.2.1 Modeling assumptions	44
3.2.2 Notations	45
3.2.3 Model formulation	47
3.2.4 Decision support tool	50
3.3 Computational studies	51
3.3.1 Experiment setting	52
3.3.2 Analysis on operational costs	52
3.3.3 Analysis on transshipment hub	58
3.4 Summary	61

Chapter 4 A TWO-STAGE STOCHASTIC MODEL FOR EMPTY CONTAINER REPOSITIONING WITH UNCERTAINTY..... 63

4.1 Problem description	64
4.2 Problem formulation.....	66
4.2.1 Modeling assumptions.....	66
4.2.2 Notations.....	67
4.2.3 Model formulation.....	68
4.3 Methodology - Sample average approximation	70
4.4 Computational studies	72
4.4.1 The transportation network.....	72
4.4.2 Results of the sample average approximation.....	74
4.4.3 Deterministic model vs. stochastic model.....	75
4.5 Summary	77

Chapter 5 PROGRESSIVE HEDGING STRATEGY FOR STOCHASTIC EMPTY CONTAINER REPOSITIONING 79

5.1 Scenario decomposition.....	80
5.2 Progressive hedging approximation -based algorithms for sample average approximation problem	82
5.2.1 Progressive hedging approximation -based algorithm 1	82
5.2.2 Progressive hedging approximation -based algorithm 2	84
5.2.3 Computational studies	86

5.3 Progressive hedging approximation -based algorithm with sequential sampling	91
5.3.1 Sequential sampling.....	92
5.3.2 Computational studies	95
5.4 Summary	98
Chapter 6 NON-I.I.D. SAMPLING TO ENHANCE THE SAMPLE AVERAGE APPROXIMATION METHOD.....	100
6.1 Introduction.....	100
6.2 Sampling methodology	103
6.2.1 Latin hypercube sampling	104
6.2.2 Supersaturated design	106
6.2.3 The proposed sampling method - Constructing Latin hypercube design by using supersaturated design.....	107
6.3 Computational studies	109
6.4 Summary	113
Chapter 7 CONCLUSIONS AND FUTURE RESEARCH.	115
7.1 Summary of results	115
7.2 Possible future research	118
BIBLIOGRAPHY	120
APPENDICES	135
Appendix A: The explicit form of the two-stage model P2	135
Appendix B: Data generation and cost parameter of the small-scale	

case.....	138
-----------	-----

SUMMARY

Empty Container Repositioning (ECR) has become a crucial issue due to the global trade imbalance between different regions. Thus, ECR problem has received more and more attention from both academics as well as industries in recent years. This thesis focuses on the operational ECR problem from the perspective of ocean liners.

The operational ECR problem is motivated by a real situation faced by an international shipping company. Weekly decisions are made by ocean liners in order to move empty containers from import-dominated regions to export-dominated regions given fixed vessel service schedules. In this study, we formulate the ECR problem as a time space network model under rolling horizon policy to cope with the dynamically changing environment in container shipping industry. An actual scale case study is presented. Compared with a simple rule which attempts to mimic the actual operation of a shipping liner, the proposed model is promising as the operational cost could be significantly reduced. Moreover, potential transshipment hubs are able to be identified by analyzing the transshipment activities.

Interview with shipping industries reveals that weekly container shipping decisions require forecast of future demands, remaining vessels' capacities, and supply. Due to the dynamically changing environment and the low forecasting accuracy in container shipping industry, ocean liners have to deal with uncertain information in container transportation. Motivated by this challenge, the second part of our work is to extend the proposed deterministic model to a two-stage stochastic model in dealing with the uncertainties in ECR. The Sample Average Approximation

(SAA) method is applied to solve the stochastic ECR problem with a large number of scenarios. Numerical experiments are provided to show the good performance of the scenario-based method for the ECR problem with uncertainties.

The SAA problem with a prohibitively large number of scenarios is usually a large-scale problem. It is usually difficult or time consuming to solve it. In order to solve the SAA problem efficiently, we consider applying the scenario aggregation by combining the approximate solution of the individual scenario problem. Algorithms based on the progressive hedging approximation strategy are developed to solve the SAA problem with multiple scenarios. By using the decomposition methods proposed, the sub-problem of the large-scale SAA problem could be efficiently solved by commercial software. A computational experiment is offered to demonstrate the efficiency of our solution methods.

Another key issue related to the SAA method is to generate representative samples. In this study, we empirically compare the performance of SAA method for the stochastic ECR problem under the well-studied independent and identical distribution (i.i.d.) sampling and several non-i.i.d. samplings. Moreover, inspired by the idea of U design which constructs the Latin hypercube design based on orthogonal array, we propose a non-i.i.d. sampling which takes the advantages of both Latin hypercube design and supersaturated design. Based on the supersaturated design, we can get better solutions by using the same number of scenarios. Our numerical experiments show that the SAA method for the stochastic ECR problem could be enhanced by these non-i.i.d. sampling schemes.

LIST OF TABLES

Table 3.1 The 49 ports and 44 services in consideration	52
Table 3.2 Operational costs of scenario 2	55
Table 3.3 Effects of ship capacity on transshipment activities	59
Table 3.4 Effects of cost parameter on transshipment activities.....	60
Table 4.1 The port rotation of service APX (westbound)	64
Table 4.2 Results of the SAA method (small-scale case, $N=100$, $N'=1000$, $M=20$)	74
Table 4.3 Deterministic model vs. stochastic model (small-scale case, single period)	76
Table 5.1 The performance of the progressive hedging based algorithms for small-scale cases ($L=20$, $\eta=2$).....	87
Table 5.2 Network parameters for the large-scale problem	88
Table 5.3 The performance of the progressive hedging based algorithms for large-scale cases ($L=10$, $\eta=5000$)	89
Table 5.4 Results of the SAA method (large-scale case, $N=30$, $N'=300$, $M=10$).....	89
Table 5.5 Deterministic model vs. stochastic model (large-scale case, single period)	90
Table B.1 Cost parameters of ECR problem (small-scale case)	139

LIST OF FIGURES

Figure 3.1 General decision processes for empty container repositioning	39
Figure 3.2 The time space network with an inter-region service and an intra-region service	42
Figure 3.3 Operational costs under different lengths of planning horizon	54
Figure 3.4 The simple rule model vs. scenario 2 of our model	56
Figure 3.5 Sensitivity analysis on cost parameters	57
Figure 4.1 The two-stage time space network	65
Figure 4.2 A network with three services and five ports.....	73
Figure 4.3 Improvement of $\hat{z}_N(\hat{x}_N^j)$ with the increase of sample size N	74
Figure 4.4 Weekly cost of the stochastic model and deterministic model overtime....	76
Figure 5.1 PHA-based algorithm with sequential sampling	94
Figure 5.2 The estimated actual cost and the best objective value (case 1).....	96
Figure 5.3 The convergence of the PHA-based algorithm with sequential sampling (case 1, 10 replications)	97
Figure 5.4 The estimated actual cost and the best objective value (case 2).....	97
Figure 5.5 The convergence of the PHA-based algorithm with sequential sampling (case 2, 10 replications)	97
Figure 5.6 The estimated actual cost and the best objective value (case 3).....	98
Figure 5.7 The convergence of the PHA-based algorithm with sequential sampling (case 3, 10 replications)	98
Figure 6.1 Probability plot of the actual cost estimates	111
Figure 6.2 Box plot of the actual cost estimates	111

LIST OF SYMBOLS

V	The set of services;
P	The set of ports in the target region;
Q	The set of regions;
K	The set of container types;
S_v	The set of stops on service v , which is defined to represent the port sequence of service v ;
$A_{v,s}$	The set of periods in which service v arrives at its stop s ;
$D_{v,s}$	The set of periods in which service v departs from its stop s ;
T	The length of planning horizon;
$c_{t,i,k}^u$	Cost of unloading an empty container of type k from a ship at port i at time t ;
$c_{t,i,k}^w$	Cost of loading an empty container of type k to a ship at port i at time t ;
$c_{t,s,v,k}^x$	The transportation cost for an empty container of type k leaving the stop s which is on service v at time t ;
$c_{t,i,k}^y$	Daily cost for storing an empty container of type k at port i at time t ;
$c_{t,i,k}^z$	Penalty cost when demand of empty container of type k in port i cannot be satisfied by the inventory at port i ;
$b_{v,s}$	Transportation time from stop s to next stop on the service v ;
$d_{v,s}$	The number of days that the service v stays at stop s ;
$\gamma_{t,s,v}$	Residual space capacity on service v when it leaves stop s at time t ;
$\sigma_{t,s,v}$	Residual weight capacity on service v when it leaves stop s at time t ;

- g_k Volume of one container of type k ;
- h_k Weight of one container of type k ;
- $\theta_{t,i,k}$ The quantity of the supply of empty containers of type k in port i at time t ;
- $\psi_{t,i,k}$ The quantity of the demand of empty containers of type k in port i at time t ;
- $p_{v,s}$ The port corresponding to the stop s on service v ;
- $f_{t,i,k}$ Total number of empty containers of type k that should be repositioned from the target region to region i at time t ($i \in Q$);
- $u_{t,s,v,k}$ The number of empty containers of type k unloaded at stop s on service v at time t ($k \in K, v \in V, s \in S_v, t \in A_{v,s}$);
- $w_{t,s,v,k}$ The number of empty containers of type k loaded from stop s onto service v at time t ($k \in K, v \in V, s \in S_v, t \in D_{v,s}$);
- $x_{t,s,v,k}$ The number of empty containers of type k transported from stop s to next stop on service v leaving stop s at time t ($k \in K, v \in V, s \in S_v, t \in D_{v,s}$);
- $y_{t,i,k}$ The number of empty containers of type k stored at port i at time t ($i \in P, k \in K, t = 1, 2, \dots, T$);
- $z_{t,i,k}$ The demands of type k container that cannot be satisfied by the existing or repositioning empty container inventory at port i at time t ($i \in P, k \in K, t = 1, 2, \dots, T$);
- Ω The set of all possible scenarios;
- ω A scenario that is unknown when decisions at stage 1 are made, but that is known when the decisions at stage 2 are made ($\omega \in \Omega$);
- $\xi(\omega)$ Parameters of the uncertain variables (demand, supply, residual ship weight capacity and residual ship space capacity) in scenario ω ;

- v The vector of ending container states of stage 1. It is the empty container inventory at each port and at each vessel at the end of stage 1 (the number of elements in v is R , $v = \{v_1, v_2, \dots, v_R\}$);
- z^* The optimal value of the true problem;
- \hat{z}_N The optimal value of the SAA problem;
- $E(\hat{z}_N)$ The expected optimal value of the approximated problem, which is also the expected perceived cost;
- $\hat{z}_N(\hat{x}_N^j)$ The actual cost estimate corresponding to the candidate solution of \hat{x}_N^j ;
- \hat{x}^* The optimal solution of the SAA problem which provides the smallest estimated actual objective value;
- λ The Lagrangian multipliers;
- ρ The penalty ratio for the differences between the scenario solutions and the overall solution;
- \bar{v} The overall solution and the reference point;
- LB The lower bound;
- UB The upper bound;
- G The estimate of the differences between the scenario solutions obtained and the reference point.

Chapter 1 INTRODUCTION

Containerization has become more and more popular in global freight transportation activities, especially in international trade routes since 1970s. Containerization helps to improve port handling efficiency, reduce handling costs, and increase trade flows. In 2004, over 60% of the world's maritime cargos were transported in containers, while some routes among economically strong countries were containerized up to 100% (Steenken et al., 2004). According to Rodrigue et al. (2009), empty containers account for about 10% of existing container assets and 20.5% of global port handling. One main issue in containerized transportation is the imbalanced container flow, which is the result of imbalanced global trade between different regions. Under this imbalanced situation, empty containers have to be repositioned from export-dominated ports which need a large number of empty containers to import-dominated ports which hold a large number of surplus empty containers. The operational cost spent on repositioning empty containers increases along with the global containerization. It is reported that empty containers have accounted for at least 20% of global handling activity since 1998 (Drewry Shipping Consultants, 2006/07). Thus, maintaining higher operational cost efficiencies in repositioning empty containers becomes a crucial issue.

To reposition containers from import-dominated regions to export-dominated regions, maritime transportation plays an important role because of its low cost and high capacity. As one of the parties operating maritime transportation, ocean liners

which manage a fleet of vessels and a large number of containers have to make Empty Container Repositioning (ECR) decisions at different levels. At operational level, short-term decisions are made by ocean liners in real-time operation. These operational decisions focus on when and how many empty containers should be moved from import-dominated ports to export-dominated ports in order to meet customer demands while reducing operational costs. However, ocean liners face some challenges while making operational ECR decisions. Firstly, the complexity in the typical ocean transportation network has made the ECR operation time consuming and difficult to conduct. Due to this difficulty, the managers of ocean liners adopt a hierarchical and sequential method to make ECR decisions, and such a method may cause cost-inefficiency. In order to reduce the inefficiency in current operation, Feng and Chang (2008) tried to apply optimization techniques to the real-scale ECR problem. Another challenge is that ocean liners have to deal with some uncertain factors like the actual transportation time between two ports/deports, the demand and supply in the future, the in-transit time of returning empty containers from customers, and the available capacity in vessels for empty containers transportation, etc. Given some of these uncertain factors in maritime transportation, Francesco et al. (2009) proposed a multi-scenario model to address the ECR problem in a scheduled maritime system.

In the subsequent section, we first provide an overview of the current operation of ECR in shipping industry and the ECR problem with uncertainties. The research scope and objective of this thesis is then described in Section 1.2. The organization

of this thesis is given in Section 1.3. A more detailed discussion of previous and on-going research will be presented in Chapter 2.

1.1 Background

ECR problem is a widely considered issue by international transportation companies, container terminals, and container leasing vendors, etc. In this section, we will provide some background information of ECR problem on the perspective of shipping companies.

1.1.1 Overview of empty container repositioning operation in shipping industry

Shipping companies provide transportation service by operating a fleet of vessels. Their container vessels transport containers from one sea port to another sea port along regular long-distance maritime routes according to a published schedule of sailing. Besides vessels, shipping company usually owns an inventory of containers to load cargos. In order to increase the utilization of containers, containers need to be loaded with cargos for a new destination as soon as possible after being emptied from cargos. However, this is not always possible due to the trade imbalance between different regions and this has resulted in holding large inventory of empty containers by ocean liners and thereby increasing the operating cost.

The physical shipping network is composed of ports, container vessels, and links between ports. For an international shipping company, its transportation service

usually covers several continents and thus the transportation network is complex and large. In addition, shipping companies have to deal with the dynamic environment while making real-time operation as related information, e.g., empty containers that returned by customers, empty containers that picked up by customers, and vessel capacity, is updated time by time. Due to the complexity in the network and the dynamic environment, the ECR operation is time-consuming and difficult to conduct. To deal with these difficulties, the managers of shipping companies adopt a hierarchical and sequential method to make ECR decisions. The global network is decomposed into several regions and vessels are considered one by one sequentially to do ECR. However, this hierarchical and sequential method may lead to inefficient decisions.

Although there are some substantial studies applying optimization techniques to the ECR problems, e.g., Feng and Chang (2008) developed a two-stage model to deal with the ECR problem involving 17 services for intra-Asia transportation, we find that the existing studies are inadequate in addressing the actual scale ECR for ocean liners at real-time operation. One limitation of most existing literature on ECR problem is that shipping practices, e.g. actual ship schedule, real scale of the network, transpiration constraints, are not well considered, and thus these studies are difficult to implement in shipping industry. Moreover, most existing works focus on analyzing the operational cost and empty container inventory at ports. Transshipment activities of empty containers have not been considered yet. Therefore, there is a need to study the ECR problem which takes into account the realistic constraints as

well as the transshipment activities related to ECR.

1.1.2 Uncertainties in maritime empty container repositioning

Due to the long transportation time of the maritime ECR, a shipping company has to make ECR decisions based on forecasting for unrealized information. Some forecasting has high accuracy, e.g., because of the booking system used in the maritime transportation, demand, supply and ship available capacity in the near future (within one week) could be forecasted accurately. This forecasting could be considered as deterministic information. However, it is difficult to obtain accurate forecasting for other information, e.g., container demand and supply more than one or two weeks. These inaccuracies in forecasting lead to the uncertainties in ECR. In the maritime transportation, container operators have to deal with a number of uncertain factors like the real transportation time between two ports/deports, future demand and supply, the in-transit time of returning empty container from customers, and the available capacity in vessels for empty containers transportation, etc. In the current shipping industry, container operators make decisions based on the nominal forecast value. Because of the differences between the expected value and the realized value, inefficient solutions may be produced.

To solve the ECR problem with uncertainties is challenging. To incorporate uncertain parameters, stochastic programming is developed to describe the ECR problem with uncertainties. Furthermore, the stochastic programming for ECR is

difficult to solve as it is difficult to estimate the operational cost under uncertainties. Advanced techniques have to be developed to solve the stochastic ECR problem efficiently. Our study on ECR problem with uncertainties is motivated in dealing with these difficulties.

1.2 Research scope and objectives

This thesis studies operational ECR problem. There are two research gaps for the ECR problem. Firstly, the existing studies are inadequate in addressing the actual scale ECR for ocean liners at real-time operation. In particular, transshipment activities of empty containers have not been considered yet. The second gap is that no existing studies address the stochastic ECR problem with a large number of scenarios where the distribution of uncertain parameters can be estimated through historical data.

The main aim of this thesis is to apply the optimization techniques to the real-time empty container operation. The specific objectives of this thesis are to:

- Develop a deterministic time space network model for ECR, where the real scale maritime transportation network and actual services are taken into account. Based on this model, both the operational cost for ECR and the transshipment activities related to ECR are analyzed.

- Propose a stochastic model which is developed based on our deterministic model to incorporate uncertainties and solve this stochastic model by applying the Sample Average Approximation (SAA) method.
- Develop scenario decomposition algorithms based on the progressive hedging approximation strategy to solve the large-scale SAA problems.
- Propose and analyze more representative sampling schemes to enhance the performance of the SAA method.

The results of our study may be significant for several reasons:

- The optimization model could be easily applied to the shipping industry as our model considers the actual service schedule and most port requirements.
- The operational cost for ECR may be reduced by applying this optimization technique. It could provide some evidences on the potential transshipment hubs for ECR by analyzing the transshipment activities of empty container.
- The stochastic model which considers some uncertain parameters may provide more robust decisions, and thus the operation cost for ECR may be further reduced.
- The progressive hedging method developed to solve our SAA problem for the ECR problem could be easily applied to solve other stochastic programs which consider a large number of scenarios.
- The performance of the SAA method could be enhanced by well-planned samplings, and thus better solutions may be obtained by solving SAA problem with the same number of sample scenarios.

- The proposed sampling designs could be applied to systems involving a large number of random variables and each experiment of the system is complex and time-consuming.

The focus of this thesis is to make maritime ECR decision for shipping companies. We only consider one transportation mode, i.e., by vessels. Other transportation modes like by rail, by truck, and by barge are not considered in this study. To simplify the problem, some assumptions are made in this study. Firstly, container substitution is not considered in this study as container substitution does not frequently happen in shipping industry (less than 20%). Secondly, we assume that service schedule is given and fixed in the planning horizon. This assumption is valid as the planning horizon of our operation model is short (several weeks), and the service schedule is not changed frequently. Note that we do not make decisions on laden container transportation in this study. As laden container transportation problem and ECR are usually considered separately in current shipping industry, and laden container has higher priority, our model is to make ECR decisions after the laden container transportation is planned.

1.3 Organization of thesis

The thesis consists of seven chapters. The rest of this thesis is organized as follows.

Chapter 2 introduces existing studies on the ECR problem.

In Chapter 3, the general decision process of making ECR decisions adopted by

shipping companies is described and a deterministic model based on the time space network is developed to formulate the problem. The actual operations and constraints of the problems faced by the liner operator are considered. A real scale case study which considers 49 ports and 44 services is presented.

In order to incorporate uncertainties in the operations model, we formulate a two-stage stochastic programming model considering random demand, supply, ship weight capacity and ship space capacity in Chapter 4. To solve the stochastic programs with a prohibitively large number of scenarios, the SAA method is applied to approximate the expected value function.

Chapter 5 presents algorithms based on the progressive hedging approximation strategy to solve the large-scale SAA problem with a large number of scenarios. Scenario aggregation is applied by combining the approximate solution of the individual scenario problems.

In Chapter 6, the performance of SAA method for the stochastic ECR problem under several non-independent and identically distributed sampling schemes is analyzed. Moreover, inspired by the idea of U design which constructs the Latin hypercube design based on orthogonal array, we propose a non-i.i.d. sampling which constructs the Latin hypercube design based on a supersaturated design.

The final chapter, Chapter 7, concludes this thesis and presents several directions for future research.

Chapter 2 LITERATURE REVIEW

This chapter presents a survey of literature pertinent to studies on Empty Container Repositioning (ECR) problem in Section 2.1. In addition to the review on ECR operations, previous and on-going studies on how to solve the stochastic ECR problem are discussed and the potential drawbacks of the state-of-art extraction method are evaluated to highlight the rationale for the alternative method proposed in the present study.

2.1 Empty container repositioning problem

Since 1970s, studies considering empty container flow management increased steadily. Generally, these studies could be classified into three levels according to the planning horizon of decisions, i.e., strategic level, tactical level, and operational level. In this section, we review studies at these three levels respectively in 2.1.1-2.1.3. Among these studies, literature on empty container operation under uncertainties is separately presented in 2.1.4.

2.1.1 Strategic level empty container repositioning

Strategic level problems of ECR are to make long-term decisions (usually longer than one year) with empty container flow in consideration. One direction of the strategic level studies pays attention on price strategy where allows realized demands to be affected by pricing. Gorman (2002) proposed a freight carrier's pricing strategy

in a network, where equipment repositioning was considered if the demand flow in the network was unbalanced. A subsequent study by Topaloglu and Powell (2007) provided a tractable algorithm to coordinate the pricing and fleet management decisions of a freight carrier where the cost of empty equipment repositioning played a significant role. More recently, Zhou and Lee (2009) studied the price strategy and competition of two transportation companies. Their studies for the first time analyzed the prices optimization and the outcome of competition in a transportation market with empty equipment repositioning.

Another direction of strategic level problems with empty container flow is to study the container-sharing and route-coordination strategy. Song and Carter (2009) indentified critical factors that impact empty container movements, and evaluated four strategies of ECR among shipping companies, i.e., container-sharing with route-coordination, container-sharing without route-coordination, route-coordination without container-sharing, and neither route-coordination nor container-sharing. This study is highly commendable for providing important information on equipment and service sharing among shipping companies.

The logistic design of container liner shipping which takes ECR into account has also been studied. Imai et al. (2009) analyzed two typical service networks with different ship size: multi-port calling by conventional ship size and hub-and-spoke by mega-ship. In their study, the problem was studied in two phases: the service network design, and container distribution. Their work provided the important insight that multi-port calling is more cost-efficient under most situations.

2.1.2 Tactical level empty container repositioning

Tactical level problems of ECR are to make mid-term decisions (usually from several months to one year) with empty container flow in consideration. At tactical level, empty container flow was mainly considered in the formulation of service network design problem, ship deployment problem, fleet sizing problem, the threshold policies for empty container inventory control problem, and the policy for empty container transportation, etc.

2.1.2.1 Service network design problem

Shintani et al. (2007) formulated a two-stage model to address the design of container liner shipping service networks by explicitly taking ECR into account. The first stage model was to construct the calling port sequence. The second stage model was to estimate the profit of container management with ECR given a set of calling port sequence. A genetic algorithm-based heuristic was also developed to get the optimal port sequence. Subsequently, Chen and Zeng (2010) decomposed the first stage model of Shintani et al. (2007) into two stages. The optimization problem of container shipping network was formulated as mixed integer non-linear programming at three stages. The first stage was to get a set of calling ports given a set of candidate ports. The second stage was to construct an optimal calling sequence given a set of ports. The third stage was to determine and arrange the optimal configuration of container.

2.1.2.2 Ship deployment problem

Different from the service network design problem which is to construct transportation network and calling sequence, the ship deployment problem is to assign a fleet of ships to a given network with fixed calling sequence. Ye et al. (2007) developed a tactical model which considered container flow management and ship deployment jointly. The objective of this model was to find the optimal service frequency and the optimal traffic flow, including both empty container flow as well as laden container flow.

2.1.2.3 Fleet sizing problem

Fleet sizing problem with empty container flow in consideration also attracts more and more attention recently. Lai et al. (1995) developed a simulation model to allocate empty containers which were transported from the Middle East to ports in the Far East. The main aim of this simulation was to determine the mix of container types that the company should maintain in the long run. Safety stock and allocation policy at each Far East port were also considered in this study. This study is a major milestone in the development of simulation model for container fleet sizing problem with inventory policies. To analyze the optimal container fleet sizing under other inventory policies, Dong and Song (2009) developed a simulation-based optimization tool to optimize the container fleet sizing and the parameterized ECR policy, i.e., the two-level threshold policy, jointly. As inland container movements are usually out of the control of shipping lines and it is one of the key factors related to fleet sizing problem, Dong and Song (2012) studied how the inland transport

times and their variability affect the container fleet sizing. A simulation-based optimization model was formulated for the container fleet sizing problem in liner services with uncertain customer demands and stochastic inland transport times. Apart from the simulation model, the fleet sizing problem with ECR also has been studied analytically. Du and Hall (1997) analyzed fleet sizing problem from inventory theory and developed stock control policies for empty equipment. In their study, the stochastic processes were analytically modeled for hub-and spoke network and then compared the analytical results to Monte Carlo simulations.

2.1.2.4 Threshold policies for empty container inventory control problem

Repositioning empty containers based on inventory control policy is another topic raised in recent year, since the inventory policy is easy to understand for container operators and is easy to operate in practice. Li et al. (2004) formulated the empty container management problem in one port as an inventory problem with positive and negative demands at the same time. Their study showed that there exists an optimal policy, (U, D) , i.e., importing empty containers up to U when the empty inventory in the port is less than U , or exporting the empty container down to D when the empty inventory in the port is more than D , doing nothing otherwise. To apply the threshold policy in a more general network, Li et al. (2007) adapted this threshold policy to multi-port case. A heuristic algorithm was developed to show how to allocate the empty containers to reduce the average cost. The threshold inventory container policy also has been analyzed under other network systems.

Song (2007) studied the optimal stationary policy for a periodic-review shuttle service system with finite reposition capacity. The threshold policy was characterized by using Markov decision process approach. A subsequently study of Song and Dong (2008) applied a three-phase threshold control policy to repositioning empty container in cyclic route. The threshold values were arbitrarily determined by the average demands and the variance of the demands. A simulation model was developed to evaluate the performance of the three-phase threshold policies for ECR.

One of the fleet management problems which are fairly closed to the ECR problem is the empty vehicle redistribution problem. The threshold inventory policies for the empty vehicle redistribution problem also have been studied in recent years, e.g., Song (2005) and Song and Earl (2008) analyzed the threshold policy in two-depot service systems, and Song and Carter (2008) studied the optimal threshold policy for the hub-and-spoke transportation systems.

2.1.2.5 Policy for ECR transportation

Empty containers are transported under certain rules in shipping industry. In some cases, the destination of the empty container is determined when the empty containers are sent to a vessel from its original port. This rule could be formulated as the typical transportation model with original-destination (o-d) pair. In other cases, however, ports of destination are not determined in advance and empty containers are unloaded from vessels as needs, whereas the direction of empty container flows

is specified. Song and Dong (2011) studied the ECR policy with flexible destination ports by developing a simulation model. And numerical results showed that the new policy outperforms the conventional policy significantly in situations where trade demands are imbalanced and container fleet sizes are within reasonable range.

Based on the review of both strategic level problems and tactical level problems, we find that although empty container flow has been taken into account in different problems, the detailed empty container operations and decision-making are not considered. In the next section, we present a review focused on operational ECR problem which is faced by container operators when making daily or weekly decisions. On the other hand, based on previous studies, we also find that the transportation policy with flexible destination is highly promising, not only because its good performance under some conditions (Song and Dong, 2011), but also because it is a widely applied policy in current shipping industry. In this thesis, we aim to develop operational models under the transportation policy with flexible destination to solve the ECR problem faced by container operators.

2.1.3 Operational level empty container repositioning

Operational level ECR is to make short-term decisions (daily or weekly) for ECR operations. In this section, we review operational studies on inland ECR, maritime ECR, and the intermodal models which take both inland ECR and maritime ECR into account.

Studies on inland ECR mainly analyze empty containers transportation among container terminals, inland depots, and customers' places. Crainic et al. (1993) described the empty container management problem in land transportation and identify its basic structure and main characteristics. Dynamic and stochastic models were developed for the allocation of empty containers. This study paved the way for inland ECR; however, no numerical studies were presented in this paper. To fill this gap, several subsequent works applied the ECR model to solve practical cases. Choong et al. (2002) studied the ECR problem in Mississippi River basin area. Three types of transportation modes, i.e., barge, truck, and rail, were considered in this study. Another novel study by Julia et al. (2006) studied empty containers movements in the Los Angeles and Long Beach port area. One maritime terminal and several inland depots were considered in their model, and the results showed that cost and traffic congestion could be reduce by considering reuse of empty container. Soon after Julia et al. (2006)' work, Chang et al. (2008) also studied the empty container movements in Los Angeles and Long Beach port area by local trucks, while their study focused on container type mismatch. An optimization model was developed for the multi-commodity empty container substitution problem. As the ECR problem is closely related to the full container transportation, a decision support system was proposed by Bandeira et al. (2009) for integrated distribution of empty and full containers among customers, leasing companies, harbors, and warehouses. Their mathematical model was formulated in two stages. The first stage was to adjusting full container demands according to the empty container inventory. The second stage

model was to optimize the cost of empty and full container transportation given full containers demands. To develop more practical model, time window of the demand was taken into consideration in Zhang et al. (2009, 2010). The truck scheduling for inland container was considered in their study. A reactive tabu search algorithm and a heuristic-based algorithm were developed to solve the container truck transportation problem.

Another direction is to analyze the container transportation in maritime shipping network. The maritime shipping network usually is consisted of a group of maritime terminals, and these maritime terminals are connected by a fleet of vessels which follow a published travelling schedule. The general maritime network model for ECR was proposed by Cheung and Chen (1998). They developed a time space network model and considered main port requirements and service network in their model. Their study paved the way for maritime ECR network modeling. Moon et al. (2010) developed a mix-integer model to describe the operational ECR problem with purchasing and short-term leasing considerations. A hybrid genetic algorithm was proposed to solve their model. Multiple ports were generated randomly and considered while the real scale network and service schedule are not considered in this study. In order to apply the general networking techniques to the shipping industry, researchers tend to consider the actual services and the real scale network in the latest decade. Actual service schedule was considered in Lam et al. (2007). They developed an approximate dynamic programming approach in deriving operational strategies for the relocation of empty containers, in which both two-port

two-voyage system and multiple-port multiple-voyage system were considered. Although actual service schedule was considered in their study, the proposed dynamic approximation programming was limited to a small-scale problem as the accuracy of the linear approximation method for multi-port system was not satisfying. One paper that considered a relatively large shipping network is Feng and Chang (2008). They developed a two-stage model to deal with the ECR problem involving 17 services for intra-Asia transportation. The first stage was to estimate the empty container stock at each port and the second stage modeled the ECR planning with shipping service network. Feng and Chang's work successfully developed a real scale network for an ocean liner and analyzed port container inventory. However, real time decisions were not provided because their monthly-based model did not consider service capacity constraints and detailed services schedules.

There are also some researches considering the both the inland ECR and maritime ECR. In an early work of Shen and Khoong (1995), a decision support system was developed to solve a large-scale planning ECR problem. The decision support system deployment framework was consisted of three levels of planning, i.e., terminal planning, intra-regional planning, and inter-regional planning. However, no numerical study was presented to describe the application of decision support system in their paper. To solve the ECR problem with inland transportation and maritime transportation, more than one transportation methods are usually used to move the empty container from its origin to destination. In this case, intermodal models with multiple types of transport modes have been proposed in recent year. Erera et al.

(2005) developed a large-scale multi-commodity flow problem on the perspective of tank container operators. Multiple transportation modes including scheduled service provided by ocean carriers and unscheduled service provided by long-haul trucking, rail service, and barge feeder were considered in their study. Olivo et al. (2005) also considered the ECR problem with multiple transportation modes for logistic companies. In their study, the ECR problem was modeled on an hourly basis, while most other studies focused on a daily or a weekly basis. Modeling on an hourly basis is effective in dealing with the dynamic environment in real-time operation, although it increases the scale of network and thus increases the difficulty to solve the model.

In this section we have reviewed inland ECR problem, maritime ECR problem and the intermodal models considering both inland and maritime ECR. These prior studies enable us to have a better understanding of the general transportation for ECR. However, based on the literature we find that the current studies are inadequate in addressing the actual scale ECR for ocean liners at real-time operation. Moreover, most current works focus on analyzing the operational cost and port empty container inventory. Transshipment activities of empty containers have not been considered yet. In view of the above issues, further research on the real scale ECR for maritime container operation is imperative. To fill this gap, one aim of our study is to develop a mathematical model for ECR under the transportations policy with flexible destination, where the real scale maritime transportation network and actual services are taken into account. Based on this model, both the operational cost for ECR and the transshipment activities related to ECR are analyzed.

2.1.4 Empty container repositioning with uncertainty

Studies reviewed in previous sections (2.1.1-2.1.3) mainly consider deterministic ECR problems, in which all information is given. However, In the maritime transportation, container operators have to deal with some uncertain factors like the real transportation time between two ports/deports, future demand and supply, the in-transit time of returning empty container from customers, and the available capacity in vessels for empty containers transportation, etc. There are several studies taking the uncertain nature of parameters into account. In an early work of Cheung and Chen (1998), a two-stage stochastic network model was developed to determine the maritime ECR and leasing decisions. All information in the first stage was given while some parameters in the second stage were uncertain when decisions in the first stage were made. The two-stage modeling is highly significant in that it successfully combines the deterministic information and the uncertain information in the ECR. A stochastic quasi-gradient method and a stochastic hybrid approximation procedure were applied to solve their stochastic model. The performance of these solving methods mostly depends on the selected approximation function. Another powerful technique for solving dynamic, stochastic programs is approximate dynamic programming, which has been widely used to solve stochastic fleet management problem. The approximate dynamic programming method has been applied to solve the ECR problem in maritime shipping network. In Lam et al. (2007)'s study, the ECR problem was formulated as a dynamic stochastic programming with the decision policy optimal in the infinite horizon average cost sense. Linear

approximation architecture was chosen to approximate the cost function. They showed that linear approximation may be insufficient to fully describe the cost function for the multi-port multi-service system. While approximate dynamic programming is a good approach to solve the ECR problem, one of the main challenges in approximate dynamic programming is the identifying a good cost to go function to represent the actual future cost. Innovative modeling which is able to exploit the structure of the problem is necessary for the function to have sufficient accuracy. Erera et al. (2009) presented a robust optimization framework based on time space network for dynamic ECR problems. The robust repositioning plan was developed based on the nominal forecast value and could be adjusted under a set of recovery sections. The advantage of this approach is that it is consistent with the current ECR operation and easy to apply. Francesco et al. (2009) proposed a multi-scenario model to address the ECR problem in a scheduled maritime system. This scenario-based model is promising because deterministic optimization techniques could be applied to solve the stochastic ECR problem. By considering more information on the uncertain parameters, this scenario-based method can provide better ECR decisions than the current approach in shipping industry which only considers the expected value of the uncertain parameters. In their study, opinions of shipping companies were considered to generate scenarios when the distributions of uncertain parameters cannot be estimated through historical data.

In summary, there are generally three types of approaches to address the stochastic maritime ECR problem. One is to consider the expected value of the

uncertain parameters, and adjust decisions according to a set of recovery sections. Although this approach is consistent with the current ECR operation and easy to apply, uncertain information is not considered in the original decision making, inefficiencies may be caused by adjusting the decisions. The second type of method is to use an approximation cost-to-go function to describe the cost in the future. Despite many successful applications, the performance of this method depends on how the cost-to-go function is approximated when applying it to a practical case. The third type of method is to develop a scenario-based model. The multiple-scenario model provided by Francesco et al. (2009) is subject to a small number of scenarios. The advantage of this method is that there is no need to approximate the value function. Moreover, deterministic optimization techniques could be applied to solve the stochastic ECR problem. In shipping industry, ocean liners usually keeps historical data on some uncertain parameters. Based on these data, distributions of these parameters are able to be estimated. Random scenarios could be generated based on these distributions. To our knowledge, no existing studies address the stochastic ECR problem with a large number of scenarios where the distribution of uncertain parameters can be estimated through historical data. To fill in this gap, the second aim of this thesis is to propose a stochastic model which incorporates uncertainties based on our deterministic model and then solves this stochastic model by applying scenario-based method, like the Sample Average Approximation (SAA) method.

2.2 Methods to solve stochastic empty container repositioning problem

As discussed in Section 2.1, further study on operational ECR problem with uncertainties in maritime shipping is imperative. Since the stochastic fleet management model is usually difficult to solve, plenty of approaches have been proposed to solve it. In this section, we first present some general methods to solve the stochastic fleet management problem in Section 2.2.1. Among these general methods, the SAA method which is a widely used scenario-based method is reviewed in Section 2.2.2. As the SAA problem with multiple scenarios is usually has large scale. Innovative approaches are needed to deal with SAA problem with multiple scenarios. Studies related to solve the SAA are mainly focused on two directions. Section 2.2.3 presents scenario decomposition approaches to decompose the large-scale SAA problem into a group of sub-problems. Finally, Section 2.2.4 reviews studies related to using more representative samplings to enhance the performance of the SAA method.

2.2.1 General methods for stochastic fleet management problem

The stochastic fleet management problem is difficult to solve in a dynamic environment. Several general stochastic programming approaches are applied to solve the stochastic fleet management with a large number of random variables and with a network structure. There are a number of approaches to solve the optimization problem with uncertainties, the most important of which could be classified into

three directions (Powell, 2009) and are introduced respectively as follows.

- Simulation optimization - According to the myopic policies, the decision function depends only on what we know now, and makes no attempt to use any sort of forecast of how decisions now might impact the future. These policies often depend on a set of parameters. Thus, we want to find the set of parameters that perform the best over many realizations of the future. The tools of stochastic search (Spall, 2003; Zhang et al., 2005) and the closely related field of simulation optimization (Fu, 2002; Fu et al., 2005; Chang et al., 2007) are typically used to choose the best set of parameters.

- Dynamic programming - Decisions are made by estimating the value of the states. In the value function approximations, Bellman's equation is used to approximate the value function. Decisions are made by solving

$$A^\pi(S_t) = \arg \max_a (C(S_t, a) + \gamma \mathbb{E}\{\bar{V}_{t+1}(S^M(S_t, a, W_{t+1})) | S_t\}) \quad (2.1)$$

where $A^\pi(S_t)$ is the function for action a if we are in state S . $\bar{V}(S)$ is an approximation of the value of being in state S . The space of this approach is the space of potential value function approximations (Powell, 2007). Approximate dynamic programming has been proved to be a powerful technique for solving dynamic programs, both in academic and in practice. It has been widely used to solve stochastic fleet management problem (White, 1972; Powell et al., 1995; Marar, 2002; Powell et al., 2006). For example, Powell and Topaloglu (2005) used piecewise linear approximations for freight car distribution. Topaloglu and Powell (2006) showed that the approximate dynamic programming-based method works

well for multi-commodity problem that arises in the dynamic resource allocation problem. Simão et al. (2009) successfully applied approximate dynamic programming to a large-scale case which involves over 6000 drivers at a high level of detail. Approximate dynamic programming also has been applied to solve the ECR problem in maritime shipping network. In Lam et al. (2007)'s study, the ECR problem was formulated as a dynamic stochastic programming with the decision policy optimal in the infinite horizon average cost sense. They showed that linear approximation may be insufficient to fully describe the cost function for the multi-port multi-service system. While approximate dynamic programming is a good approach to solve the problem, one of the main challenges in approximate dynamic programming is the identifying a good cost-to-go function to represent the actual future cost. Innovative modeling which is able to exploit the structure of the problem is necessary for the function to have sufficient accuracy.

- Stochastic rolling horizon procedures - The stochastic rolling horizon procedure use a stochastic forecast of future events based on what we know at time t . Based on the forecast, we then solve a problem that extends over a planning horizon, but only implement the decision for the immediate time period.

Interviews with industries show that the container shipping decisions are made based on forecasting. Due to the dynamically changing environment and the low forecasting accuracy in container shipping industry, the forecasting has to be adjusted when new information is updated. Considering the nature of the three approaches reviewed above and the changing environment in container shipping

industry, the stochastic model in rolling horizon procedures could be applied to deal with the dynamically changing forecasting for ECR. Besides, the forecasting for container (ship capacity, demand, and supply) would be difficult after one or two weeks. In this case, it is difficult to collect data to generate the sample path in multiple stages for our ECR problem (e.g., the probability of states s' at stage $t+1$ given the states s at stage t is difficult to get). Thus, in this thesis, we aim to develop a two-stage stochastic model in rolling horizon procedure to solve the operational ECR problem under uncertainties.

There are many methods applied to solve the two-stage stochastic fleet management problem. Stochastic linearization method was applied to solve the two-stage stochastic ECR problem in Cheung and Chen (1998). The idea of this method is to replace the nonlinear expected resource function with a sequence of weighted averages of stochastic sub-gradients that are linear (Ermoliev, 1988; Soize, 1995). The main limitation of this method is that obtaining stochastic sub-gradients can be computationally expensive because a large number of samples may be needed (Hurtado and Barbat, 1996). The Benders' decomposition (Benders, 1962; Lasdon, 2002) is also applied to solve the stochastic network programs (Cordeau et al., 2000). This technique allows the solution of vary large linear programming problems that have a special block structure. As it progresses towards a solution, Benders' decomposition adds new constraints. This method produces highly fractional solutions and may be subject to low convergence rate. Another direction is to derive separable approximations of the value function, e.g., the auxiliary function method

(Silva et al., 2003; Wu et al., 2008), the stochastic hybrid approximation procedure (Cheung and Powell, 2000), and the structured adaptive function algorithm (Sanger, 1991; Wallace and Ziemba, 2005). One challenge of applying these approaches to practical problems is that we have to solve the curses of dimensionality when deriving separable approximations of the value function. What remains is scenario-based method to solve the stochastic programs. In next section, we establish a body of literature on scenario-based methods, particularly, one widely used method, i.e., the SAA method, will be introduced.

2.2.2 Sample average approximation method

The general two-stage ECR problem could be formulated as

$$z^* = \min_{x \in X} c^T x + E_p[Q(x, \xi(\omega))] \quad (2.2)$$

The overall objective is to minimize the first stage ECR operational cost and the expected ECR cost at the second stage. $E_p[Q(x, \xi(\omega))]$ is the expected cost at stage 2, where p is the probability distribution of uncertain parameters. The exact evaluation of the expected value is difficult, i.e., $E_p[Q(x_1, \xi(\omega))]$. It requires the solutions of a large number of stage 2 optimization problems. Therefore, the problem could be unmanageable large and becomes time-consuming to solve. There are a large number of studies to address this problem. One direction is to use integration methods (Acevedo and Pistikopoulos, 1996; Acevedo and Pistikopoulos, 1998; Pintarič and Kravanja, 2000). For example, Novak and Kravanja (1999) proposed an

approximation method which used extreme points. The weighted average over critical points was considered in the objective function. Another direction is sampling method. For stochastic programs with a large number of scenarios, sampling based approaches have been proposed to estimate function values, optimality cuts, gradient, or bounds for the second-stage expected value function (Kim and Diwekar, 2002).

Generally, sampling-based approximation approaches could be classified into two main groups: internal samplings and external samplings. In internal sampling method, new samples are generated and added to previous generated samples over iterations inside an algorithm. For example, Van Slyke and Wets (1969) suggested modifying the samples based on L-shaped algorithm for stochastic linear programming. Higle and Sen (1991) proposed a stochastic decomposition algorithm which updated a piecewise linear approximation of the expected function in a sub-problem iteratively. Norkin et al. (1998a) and Norkin et al. (1998b) developed the sampling branch and bound method to solve the discrete stochastic problems. Importance sampling is considered for the stochastic large-scale linear problem in Infanger (1994).

In external samplings, sample scenarios are generated from Ω , the true problem is approximated by the SAA problem,

$$z^* = \min_{x \in X} c^T x + \frac{1}{N} \sum_{n=1}^N Q(x, \xi(\omega)) \quad (2.3)$$

The basic idea of SAA method is that the expected objective function of the stochastic problem is approximated by a sample average estimate derived from a

random sample and the resulting SAA problem could then be solved by deterministic optimization techniques (Rubinstein and Shapiro, 1990; Geyer and Thompson, 1992). In earlier work, studies on SAA method are mainly for the stochastic linear problem (Plambeck et al., 1996). The convergence rate of the SAA method and statistic inference for the stochastic linear problem was analyzed in Shapiro (1996). The solution quality of the SAA method is estimated by the optimality gap, which is the difference between the lower bound (i.e., the average objective value of multiple replications of SAA problem) and the upper bound (i.e., the objective of one candidate solution). The bounding techniques for determining solution quality in stochastic linear programs were studied in Mak et al. (1999). The behavior of the SAA method for the stochastic discrete programs was first studied in Kleywegt et al. (2002). A numerical example for the stochastic knapsack problem was also presented in their study. More recently, Wei and Realff (2004) developed two algorithms based on the SAA method to solve the convex stochastic mixed-integer nonlinear programs, and their case studies showed that the computational time required could be significantly reduced by applying their algorithms.

The applications of the SAA method to the two-stage stochastic programs were also presented in several papers. For example, the two-stage stochastic program with linear resource was first studied with SAA method in Shapiro and Homem-de-Mello (1998). Verweij et al. (2003) presented a detailed computational study which applied the SAA method to the two-stage stochastic routing problems. Decomposition and branch-and-cut were applied to approximate the problem under the SAA scheme in

their study. To analyze the quality of solutions of the SAA method for different stochastic problems empirically, Linderoth et al. (2006) developed a computer software grid to solve many large instance of two-stage stochastic linear programs. On the other hand, the two-stage stochastic program with integer recourse also studied in Ahmed and Shapiro (2002).

In view of the established literature on the SAA method above, we find that SAA method has been proved to be a powerful technique to solve the stochastic programs both in academic and in practice. Thus, SAA is a promising method to solve our two-stage stochastic ECR problem as it has been successfully applied to solve a lot of two-stage network problems. However, although these prior studies enable us to have a better understanding of the SAA method, they are not specific enough to address our ECR problems directly. To our knowledge, no existing studies apply the SAA method to the stochastic ECR problem with a large number of scenarios. This study is to fill in this gap. To successfully apply the SAA method to our ECR problem, we have to take the nature the operational ECR problem, e.g., the complexity of transportation network, and the large number of random parameters, into account carefully. These challenges highlight the importance of developing intelligent techniques in dealing with the large-scale SAA problem with multiple scenarios. In next two sections, studies related to how to efficiently solve the SAA problem with multiple scenarios are reviewed.

2.2.3 Scenario decomposition for the stochastic problem with multiple scenarios

To successfully apply sampling techniques, e.g., the SAA, to solve the stochastic ECR problem, one must be able to develop an efficient algorithm to solve the approximated stochastic ECR problem, through the use of scenarios. In this section, we establish a detailed review on algorithms based on scenario decomposition in optimization under uncertainty. The problem of identifying a set of representative scenarios will be covered in the Section 2.2.4.

The SAA problem with multiple scenarios is typically difficult to solve due to its large scale. Several decomposition algorithms which have been developed to take advantage of the special structure of the problems (Birge, 1997; Ruszczyński, 1997) is introduced in this section.

Cutting plane algorithms such as Benders decomposition (Benders, 1962) was applied to solve the large-scale SAA problem for supply chain network design under uncertainty (Santoso et al., 2005). The L-shaped decomposition method (Van Slyke and Wets, 1969) was operated in their study. While the Benders decomposition algorithm is a finite scheme, the number of iterations required may be too large in practice. Thus, techniques may have to be developed to accelerate the Benders decomposition.

Another widely used scenarios decomposition method is based on the progressive hedging strategy. In Rockafeller and Wets (1991), a rigorous algorithm procedure for scenario and policy aggregation in response to any weight of the

scenarios was proposed for the first time. Augmented Lagrangean strategy that converges to a global optimum in the case of continuous stochastic problems was suggested. Different from the node-wise decomposition (Salinger, 1997; Korf, 1998), the progressive hedging strategy is based on scenario-wise formulation, which leads to decomposition over the scenarios, and to a very simple master problem (Pennanen and Kallio, 2006). Due to this advantage of the progressive hedging algorithm, it has been widely applied to solve stochastic programming problems like fisheries management problem (Helgason and Wallace, 1991), mixed integer (0, 1) multistage stochastic programming (Løkketangen and Woodruff, 1996), stochastic lot-sizing problem (Haugen et al. 2001), stochastic inventory routing problem (Hvattum and Løkketangen, 2009), and stochastic network design problem (Crainic et al., 2011).

Based on the discussion above, we find that the progressive hedging strategy has been successfully applied to a number of stochastic problems due to its simplicity and the effective decomposition it can provide. And thus scenarios decomposition algorithm based on the progressive hedging strategy is promising to solve the large-scale SAA problem for our ECR problem. By combining the approximate solution of the individual scenario of the ECR problem, we may get overall solution with good quality. According to our knowledge, there have been no studies which apply the progressive hedging strategy to solve the two-stage stochastic ECR problem for ocean liners. And thus it will be a new contribution to the maritime community by adopting the progressive hedging strategy approximation in our ECR problem with uncertainty. Moreover, we find that current applications with the

progressive hedging algorithm are usually under given sample size. However, in review of the SAA method, we find that to identify the appropriate sample size is a crucial issue when applying the SAA method. To deal with this challenge, we aim to develop a progressive hedging based algorithm which sequentially adds new scenarios to current sample during the iteration in this study. Therefore, the sample size may not have to be determined in advance.

2.2.4 Sampling schemes to enhance the performance of sample average approximation

The SAA problem with multiple scenarios is usually difficult to solve due to its large scale. In Section 2.2.3 we have reviewed a group of scenario decomposition methods to solve the multi-scenario SAA problem. In this section, we present a review on several studies in another direction, i.e., to generate more representative samples in order to enhance the performance of the SAA method.

Independent and identical distribution (i.i.d.) sampling is well studied for construction approximations (e.g., Ahmed and Shapiro, 2002; Kleywegt et al., 2002; Ruszczyński and Shapiro, 2003, Ch. 6; Verweij et al., 2003; Wei and Realff, 2004). How to predetermine the sample size in order to estimate an approximate solution within the prescribed precision and confidence is one of the most important issues concerning the convergence analysis. The so called exponential convergence based on the Cramer's large deviation theorem (Dembo and Zeitouni, 1998) requires i.i.d. sampling. However, it is difficult or computationally expensive to obtain an i.i.d.

sampling particularly when the sample size is large. Indeed, the class of quasi-Monte Carlo method, which has become popular especially after the work of Niederreiter (1992), does not require i.i.d. sampling while works remarkably well. Under this situation, SAA under non-i.i.d. samplings received increasing attention in recent years. The convergence of SAA estimators under general sampling (including i.i.d. and non-i.i.d.), according to our knowledge, was first investigated in Dai et al. (2000). Homem-De-Mello (2008) established the exponential convergence rate under the non-i.i.d. sampling for the optimization. Some variance reduction techniques under some appropriate assumptions were theoretically studied in his study, and the results were applied to two sampling schemes, i.e., Latin hypercube sampling (McKay et al., 1979) and randomized quasi-Monte Carlo (Pennanen, 2005). More recently, SAA with general sampling was studied by Xu (2010). The uniform exponential convergence of the sample average of a class of lower semi-continuous random function was first derived in his study. Empirically, Freimer et al. (2010) investigated two sampling methods, i.e., Latin hypercube sampling and antithetic variates (Higle, 1998). The affect of these two sampling schemes on the bias and variance was computationally studied. Results of their study showed that Latin hypercube and antithetic variates methods outperformed those under i.i.d. sampling, with Latin hypercube outperforming antithetic variates.

In computer experiments, it is well known that Latin hypercube design achieves maximum stratification in one-dimensional projections. In order to possesses satisfaction in one- and higher dimensional projections, Tang (1993) used orthogonal

arrays of strength two or higher to construct Latin hypercubes, i.e., U design. Tang and Qian (2010) used U design to further enhance the accuracy of the SAA and their theoretical results showed that the SAA with U designs can significantly outperform those with Latin hypercube designs. As the U design is based on orthogonal arrays, the main problem with U design is that the required number of experiments increases exponentially with the number of parameters. In this case, it is not applicable if the system we study is hightivariate or large.

In summary, it has been shown that the non-i.i.d. sampling could help to enhance the performance of the SAA method. To our knowledge, no existing studies apply the SAA method with non-i.i.d. samplings to the stochastic ECR problem. On the other hand, although the U design has been proved to be able to further enhance the accuracy of the SAA, it could not be directly applied to our ECR problem which considers a large number of random variables. Inspired by U design, which uses orthogonal arrays to construct Latin hypercubes, we consider constructing Latin hypercubes using superaturated design, i.e., designs where the number of experiments is less than the number of variables. As a special branch of development of experiment, the construction of supersaturated design has been well studied (Booth and Cox, 1962; Lin, 1993; Lin, 1995; Nguyen, 1996; Butler et al. 2001; Bulutoglu and Cheng, 2004; Goury, 2010). Based on the supersaturated design, we may do as a whole very few experiments (even less than the number of degrees of freedom of the system when that is possible) and still get a satisfying approximation. According to our knowledge, there have been no studies focused on applying

supersaturated design to enhance SAA if the scale of the experiment is large and difficult to solve. Using non-i.i.d. samplings based on the supersaturated design to enhance the performance of the SAA method for the two-stage stochastic ECR problem will be discussed in Chapter 6.

Chapter 3 A TIME SPACE NETWORK MODEL ON EMPTY CONTAINER FLOW MANAGEMENT

Since 1970s, containerization has become increasingly popular in global freight transportation activities, especially in international trade routes. Containerization helps to improve port handling efficiency, reduce handling costs, and increase trade flows. In order to increase the utilization of containers, containers should be reloaded with new cargoes as soon as possible after reaching its destination. However, this is not always possible due to the trade imbalance between different regions in the world and this has resulted in holding large inventory of empty containers by ocean liners and thereby increasing the operating cost. Generally export-dominated ports need a large number of empty containers, while import-dominated ports hold a large number of surplus empty containers. Under this imbalanced situation, a profitable movement of a laden container usually generates an unprofitable empty container movement. The main challenge is when and how many empty containers we should move from the import-dominated ports to export-dominated ports in order to meet the customer demands while reducing the operational cost.

3.1 Problem description

From the literature in Chapter 2 we find that the current studies are inadequate in addressing the actual scale Empty Container Repositioning (ECR) for ocean liners.

Moreover, most current works focus on analyzing the operational cost and empty container inventory at ports. Transshipment activities of empty containers have not been considered yet. In order to apply the optimization techniques to ECR problem for ocean liners, we first analyze current operations of an ocean liner. Next, the time space network is introduced.

3.1.1 General decision process of empty container repositioning

The typical ocean transportation network is large and complex. Due to this complexity, the managers of shipping companies adopt a hierarchical and sequential method to make ECR decisions. This sequential method may lead to inefficient decisions. Figure 3.1 shows a real example on the sequence of decision processes of ECR for satisfying the demand in the Asia region at week t . The decisions can be divided into two phases which are described as follows.

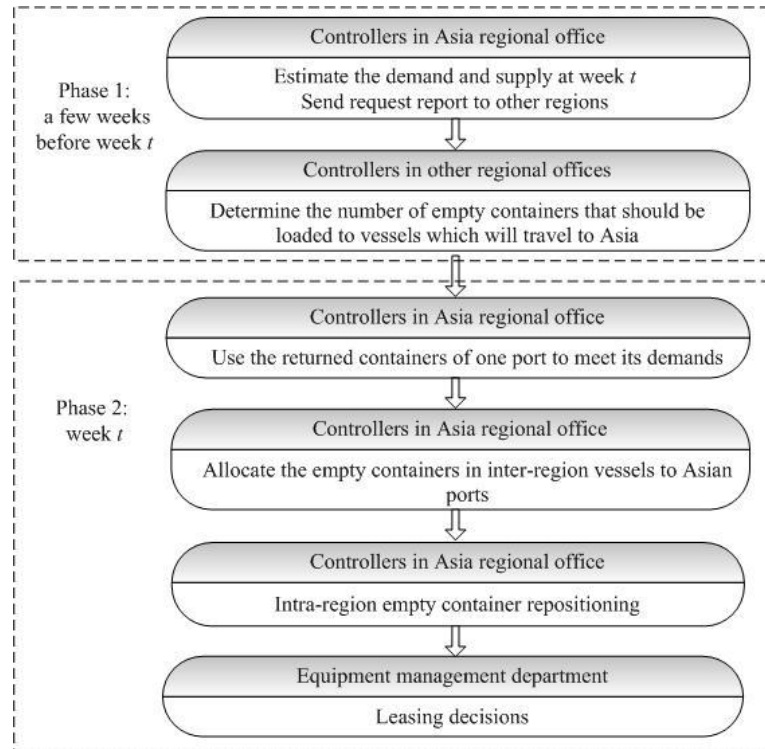


Figure 3.1 General decision processes for empty container repositioning

Phase 1: The controllers at the Asia regional office need to send requests to other regions for empty containers to be sent to the Asia region in order to meet the demand for ports in Asia at week t . Since the lead time for the inter-region transportation is usually longer than one week, the request for such inter-region empty containers is made few weeks before week t which depends on the lead time. This requires forecasting because the demand and supply (the returned empty containers) for week t during the time of request are uncertain. The forecasting takes consideration of historical demand and supply, the delinquency record and the economic environment. After receiving the request, controllers in other regional offices will determine the number of empty containers that should be loaded to vessels which will travel to Asia.

Phase 2: At the beginning of week t , the demand and supply at each port are known. Moreover, by this time, the empty containers which are loaded in the inter-region vessels (phase 1 decision) will arrive in the Asia region. Based on this information, the controller at the Asia region will use a set of simple rules to determine how to fulfill the demand. These rules are usually done in a sequential manner. The controller at the Asia region will first try to fulfill the demand for each port by using its own supply (for the export-dominated port, this supply is usually not enough). If the supply at the port is not sufficient, the controller will allocate the empty containers in the inter-region vessels to the port. If it is still not enough, the controller would then make intra-region ECR decisions, i.e., containers are transported from import-dominated ports to export-dominated ports within Asia.

Leasing decisions will be made if there are still unmet demands. Finally, when containers from the leasing vendors are unavailable or leasing decision is unprofitable, demands may be rejected by shipping liners.

3.1.2 Time space network

Container vessels are usually operated according to a published schedule of sailing and the ships often travel on closed loops. In this study, a service is defined as a specific sequence of calling ports that each ship is deployed in the loop that repeats in each voyage. Weekly service for each loop is commonly provided by most shipping companies. In this case, if the cycle travelling time of a service is four weeks, four vessels are assigned in this service. The physical shipping network is composed of ports, shipping services, and links between ports. We decompose the physical shipping network according to the geographic region. Intra-region travelling is usually completed within one week, while inter-region travelling is often much longer. In this case, when a vessel departs from the region of origin, demand and supply in the region of destination when the vessel arrives might be uncertain. Decisions on the inter-region repositioning are thus made according to forecasts. These decisions only determine how many empty containers to be loaded into the inter-region vessels, but the actual allocation of the empty containers will be done when they arrive at the region of interest. Hence, at the beginning of every week, we need to decide how to allocate the empty containers in the inter-region vessels to

various ports as well as perform intra-region ECR in order to meet the demand in each region at the lowest cost. In this study, we define the target region as the region for which the ECR decisions are made, e.g., in Figure 3.1, Asia region is the target region.

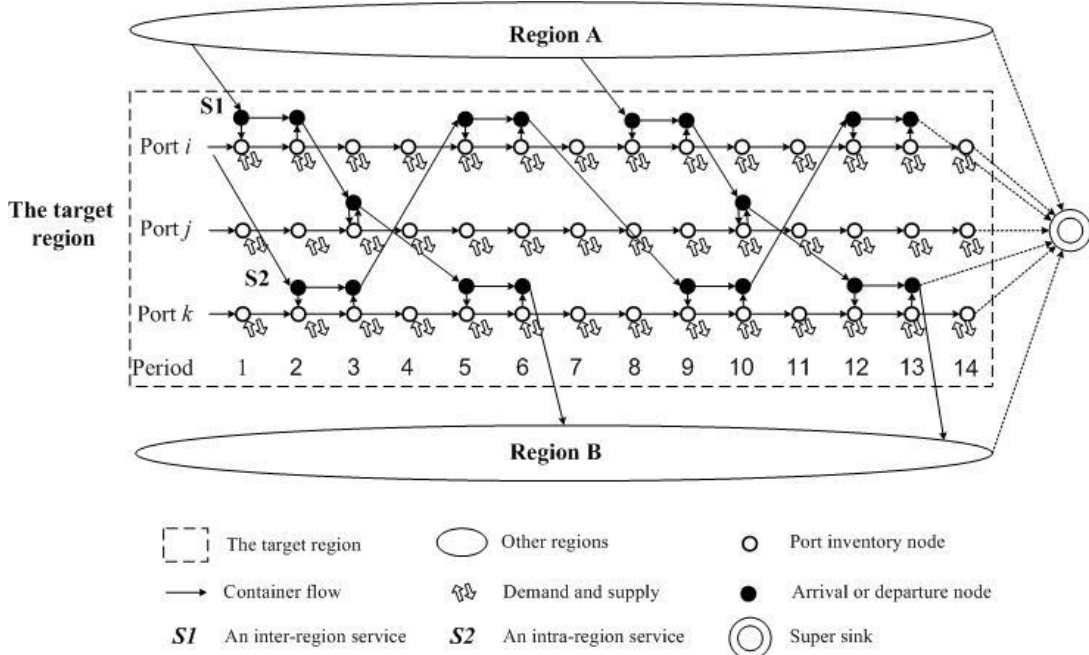


Figure 3.2 The time space network with an inter-region service and an intra-region service

We develop a time space network model to manage empty container flow, which is depicted in Figure 3.2. Two services ($S1$ and $S2$) are shown in Figure 3.2. $S1$ is an inter-region service which travels from region A to ports i , j , and k in the target region, and then travels to region B. When making decisions for the target region, all ports in a region which is not the target region (like region A and region B in Figure 3.1) are represented by one node. The number of in-transit inter-region empty containers in $S1$ from region A to the target region is known. When the inter-region services arrive, empty containers in $S1$ can be reallocated to each port in the target

region based on the updated demand (realized and forecast). $S2$ is an intra-region service which travels between port k and port i . The intra-region repositioning between port k and port i could be served by $S2$. Note that the inter-region service also could be used to do intra-region repositioning, e.g., $S1$ could reposition empty container from port i to port j .

3.2 Mathematical model

This study is based on the forecasting demand, supply, and ship residual capacity, where the supply refers to the empty containers returned by customers while the demand refers to the empty containers needed by customers to load their cargos. Residual capacity is the available ship capacity for empty containers. It is equal to the total capacity of a vessel minus the capacity of the laden containers that are being carried by that vessel. The typical transportation problem of ECR focuses on the pair of origin-destination ports. This typical formulation is easy to develop and understand. However, there are three issues when we apply this formulation to solve real scale network problems. First, due to the long transportation time and the dynamic environment, it is usually difficult to determine the destinations of empty containers in advance. In practice, empty containers are unloaded from vessels to ports according to the realized demands at ports. Second, there are a large number of decision variables and a large number of redundant constraints in the typical formulation which makes the model difficult to solve in a short time. Third, this

formulation has difficulties in dealing with multiple ships serving the same pair of origin-destination ports and leaving on the same day. This situation is not uncommon for international ocean liners' practice. In order to solve these three issues, we consider reformulating the empty container transportation problem. Other than considering the pair of origin-destination ports as in the typical formulation, we model the problem from the perspective of services. A policy with flexible destination ports is considered. In our model, redundant constraints are removed and solutions can be provided within a few seconds by commercial software so as to fulfill the frequent need of making repositioning decisions. Besides, our model takes into account multiple ships servicing the same pair of origin-destination ports which leave on the same day.

3.2.1 Modeling assumptions

The following assumptions are necessary in developing the mathematical formulation for the problem:

- (1) For the target region, the schedules of services are given and fixed;
- (2) Multiple types of empty containers are considered and they are assumed not substitutable;
- (3) Instead of considering individual demand/supply from customers, we consider the integrate demand/supply of the inland areas which are covered by each maritime terminal;
- (4) Container transportation between inland depots and maritime terminals is not

considered in this study;

- (5) When a vessel stays at a port, all unloading activities happen at the first day and all loading activities happen at the last day.

The forecasting for empty containers in the near future is more accurate due to the use of booking systems in container transportation industry. To solve this problem, we adopt a rolling horizon approach, whereas only the decisions on the current week will be adopted even though our model makes all decisions in the planning horizon.

3.2.2 Notations

Notations we adopted in proposed model are given as follows:

Index sets

- V The set of services;
- P The set of ports in the target region;
- Q The set of regions;
- K The set of container types;
- S_v The set of stops on service v , which is defined to represent the port sequence of service v ;
- $A_{v,s}$ The set of periods in which service v arrives at its stop s ;
- $D_{v,s}$ The set of periods in which service v departs from its stop s .

Parameters

- T The length of planning horizon;
- $c_{t,i,k}^u$ Cost of unloading an empty container of type k from a ship at port i at time t ;
- $c_{t,i,k}^w$ Cost of loading an empty container of type k to a ship at port i at time t ;
- $c_{t,s,v,k}^x$ The transportation cost for an empty container of type k leaving the stop s which is on service v at time t ;
- $c_{t,i,k}^y$ Daily cost for storing an empty container of type k at port i at time t ;
- $c_{t,i,k}^z$ Penalty cost when demand of empty container of type k in port i cannot be satisfied by the inventory at port i ;
- $b_{v,s}$ Transportation time from stop s to next stop on the service v ;
- $d_{v,s}$ The number of days that the service v stays at stop s ;
- $\gamma_{t,s,v}$ Residual space capacity on service v when it leaves stop s at time t ;
- $\sigma_{t,s,v}$ Residual weight capacity on service v when it leaves stop s at time t ;
- g_k Volume of one container of type k ;
- h_k Weight of one container of type k ;
- $\theta_{t,i,k}$ The quantity of the supply of empty containers of type k in port i at time t ;
- $\psi_{t,i,k}$ The quantity of the demand of empty containers of type k in port i at time t ;
- $p_{v,s}$ The port/region corresponding to the stop s on service v ;
- $f_{t,i,k}$ Total number of empty containers of type k that should be repositioned from the target region to region i at time t ($i \in Q$).

Decision variables

- $u_{t,s,v,k}$ The number of empty containers of type k unloaded at stop s on service v at time t ($k \in K, v \in V, s \in S_v, t \in A_{v,s}$);

$w_{t,s,v,k}$ The number of empty containers of type k loaded from stop s onto service v at time t ($k \in K, v \in V, s \in S_v, t \in D_{v,s}$);

$x_{t,s,v,k}$ The number of empty containers of type k transported from stop s to next stop on service v leaving stop s at time t ($k \in K, v \in V, s \in S_v, t \in D_{v,s}$);

$y_{t,i,k}$ The number of empty containers of type k stored at port i at time t ($i \in P, k \in K, t = 1, 2, \dots, T$);

$z_{t,i,k}$ The demands of type k container that cannot be satisfied by the existing or repositioning empty container inventory at port i at time t ($i \in P, k \in K, t = 1, 2, \dots, T$).

3.2.3 Model formulation

The objective function of the mathematic model (defined as **P1**) is formulated as follows,

P1:

$$\begin{aligned}
 \min \quad & \sum_{t=1,2,\dots,T} \sum_{k \in K} \sum_{i \in P} c_{t,i,k}^u \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, p_{v,s}=i, t \in A_{v,s}\}} u_{t,s,v,k} \\
 & + \sum_{t=1,2,\dots,T} \sum_{i \in P} \sum_{k \in K} (c_{t,i,k}^y y_{t,i,k} + c_{t,i,k}^z z_{t,i,k}) \\
 & + \sum_{t=1,2,\dots,T} \sum_{k \in K} \sum_{i \in P} c_{t,i,k}^w \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, p_{v,s}=i, t \in D_{v,s}\}} w_{t,s,v,k} \\
 & + \sum_{t=1,2,\dots,T} \sum_{k \in K} \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, t \in D_{v,s}\}} c_{t,s,v,k}^x x_{t,s,v,k}
 \end{aligned} \tag{3.1}$$

Considering the long travelling time, the planning horizon in our case study is assumed to be three weeks and the objective function is to minimize the total operational costs in these three weeks. The operational costs include the handling cost, the holding cost, the transportation cost, and the penalty cost. The handling cost refers to the cost of unloading empty containers from a vessel to a port and loading

empty containers from a port to a vessel. The holding cost refers to the storage cost and capital cost when empty containers are stored at ports. The transportation cost captures the cost of transporting empty containers on the vessels. The penalty cost is the cost incurred when the demand cannot be met by the existing or repositioning empty container inventory.

Ship residual space constraints

$$\sum_{k \in K} (g_k \times x_{t,s,v,k}) \leq \gamma_{t,s,v} \quad \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in D_{v,s}\} \quad (3.2)$$

Ship residual weight constraints

$$\sum_{k \in K} (h_k \times x_{t,s,v,k}) \leq \sigma_{t,s,v} \quad \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in D_{v,s}\} \quad (3.3)$$

These two constraints ensure that the total space and weight occupied by empty containers should not exceed the total available ship space and weight capacity respectively. Although the weight of empty container is light and not considered in most ECR studies, we find that it should be considered under some scenarios. For a 40' standard steel container which is used to transport general cargo, its tare weight is 3,720 kg - 3,740 kg, which is about 14% of the payload (weight 26,760 kg - 28,760 kg) of the container. If the weight capacity of a ship is almost reached, no more empty container could be loaded to the ship even if there is some space available. The left hand side of constraint (3.2) is the total space occupied by different types of empty containers. The right hand side of constraint (3.2) is the total available ship residual space capacity for empty containers. The explanation for constraint (3.3) is similar except that it is related to weight capacity. These two constraints should be considered when there is a service leaving the port.

Service flow constraints

$$x_{t-b_{v,s},s,v,k} - u_{t,s+1,v,k} + w_{t+d_{v,s+1},s+1,v,k} = x_{t+d_{v,s+1},s+1,v,k} \quad (3.4)$$

$$\forall k \in K, \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in A_{v,s+1}\}$$

$$x_{t-b_{v,s},s,v,k} \geq u_{t,s+1,v,k} \quad \forall k \in K, \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in A_{v,s+1}\} \quad (3.5)$$

Constraint (3.4) guarantees the balance of container flows at each service.

Constraint (3.5) ensures that the number of empty containers unloaded from a vessel should not exceed the total number of empty containers in the vessel. These two constraints should be considered when there is a service arriving at a port. If $t - b_{v,s} \leq 0$, $x_{t-b_{v,s},s,v,k}$ is the initial empty containers on a vessel, which is known at the beginning of the planning horizon.

Port flow constraints

$$y_{t-1,i,k} + \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s}=i, t \in A_{v,s}\}} u_{t,s,v,k} + z_{t,i,k} + \theta_{t,i,k} - \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s}=i, t \in D_{v,s}\}} w_{t,s,v,k} - \psi_{t,i,k} = y_{t,i,k} \quad \forall k \in K, \forall i \in P, \forall t = 1, 2, \dots, T \quad (3.6)$$

Constraint (3.6) considers the balance of container flows at each port at each time. The end inventory at time t is equal to the original inventory plus total container flows at time t . Note that only those services arriving or departing from the port at time t will be considered.

Container flow to other regions

$$\sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s+1}=i, t \in D_{v,s}\}} x_{t,s,v,k} = f_{t,i,k} \quad \forall t = 1, 2, \dots, T, \forall i \in Q, \forall k \in K \quad (3.7)$$

Constraint (3.7) ensures that the total number of empty containers sent to other regions should equal to the desired inter-region flow. In our model, we assume these values are known.

The non-negative constraints

$$\begin{aligned}
& u_{t,s,v,k}, w_{t,s,v,k}, x_{t,s,v,k}, y_{t,i,k}, z_{t,i,k} \geq 0 \\
& \forall k \in K, \forall i \in P, \forall t = 1, 2, \dots, T, \forall (v, s) \in \{(v, s) \mid v \in V, s \in S_v\}
\end{aligned} \tag{3.8}$$

Constraint (3.8) ensures that all variables are non-negative. According to the shipping company under study, the number of empty container movements is usually large. Therefore, the effect of using non-integer variables for number of containers is small. Fractional solutions are produced by this model, and we should perform some rounding procedures before implementing the solutions.

3.2.4 Decision support tool

In real operation planning, decision support tools are usually developed to cater for specific business requirement. We also build a decision support tool for ECR based on our mathematical model. The tool is built using ILOG Optimization Decision Manager, which is simple to use and allows users to adjust some of the assumptions, constraints, and goals for the model. This decision support tool uses scenarios to define a set of user specified data and requirements. We have considered three sets of additional rules in the decision support tool. The first scenario has an additional rule that excessive backlog or leasing quantity exceeding certain limit is not allowed. This is reasonable as excessive backlog or leasing may damage the company's reputation or incur huge extra cost. The second scenario, in addition to limiting the port shortage, each port should also keep at least a minimum level of inventory to mitigate the end of horizon effect. This is to avoid the short-sighted planning that

only meets the demand within the planning horizon without considering future demand outside the planning horizon. The third scenario takes into account the common practice of “move-count” in the industry, which limits the maximum number of handling (both loading and unloading) at a port for a given vessel. This is to cater the need to consider the port handling capacity. In the actual operation, when operators use the application, it is possible that new scenarios are added to represent other business requirements. When we are considering extra business rules, it is common that we may not be able to get a feasible solution because of the confliction between addition requirements. The decision support application is able to walk around this problem by relaxing the constraints in the case when no feasible solution is found, according to a pre-defined priority sequence. There are nine build-in priority levels: ignored, very low, low, medium low, medium, medium high, high, very high and mandatory. The application will search for relaxation from the lowest priority level to the highest one. Operations managers and planners can set the priority according to the relative importance of the business rule associated with the constraint.

3.3 Computational studies

In our numerical study, a real-life case is analyzed to show the application and performance of our model.

3.3.1 Experiment setting

Table 3.1 The 49 ports and 44 services in consideration

Ports: Bangkok, Belawan, Bugo, CagayandeOro, CatLai, Cebu, Chittagong, Chiwan, Dalian, Davao, Hakata, HoChiMinhCity, Hong Kong, Hososhima, Jakarta, Kaohsiung, Kobe, Kwan yang, Laem Chabang, Lianyungang, Manila, Moji, Nagoya, Naha, Ningbo, Oita, Osaka, Panjang, Penang, Port Klang, Pusan, Qingdao, Semarang, Shanghai, Shekou, Shibushi, Shimizu, Sihanoukville, Singapore, Subic Bay, Surabaya, Tokyo, Ube, Westport, Xiamen, Xingang, Yangshan, Yantian, Yokohama *Services: APX, ESX, JAS, NYX, PCE, PCX, PNW, PS1, PS2, PSX, SAX, AEX, CFX, ESX, EU2, EU3, JEX, MED, MEX, NCE, SCX, BKX, BLA, CHT1, CHT2, CSJ, JKT_1, JKT_2, JKT2, JSX NB, JSX SB, LCX, MAS, MDX, MNX, NVX, NKX, SEM1, SEM2, SKX, SUR1 SUR2, SVS, SVX
--

** The abbreviations of these services are listed, e.g., APX is the abbreviation of Atlantic Pacific Express*

In our numerical experiment, we consider an actual network with 49 ports and 44 services as shown in Table 3.1. These ports are mainly located in East and South Asia. Among the 44 services under consideration, there are 11 North America-Asia services, 10 Europe-Asia services, and 23 intra-Asia services. And all these services are all based on actual service schedules. These services comprise of 90% of the container flows operated by the shipping company under study. The data used in the numerical studies are based on the sample data collected and provided by the shipping company. We use CPLEX11.2 to solve the mathematical model, and the computation time is within several seconds.

3.3.2 Analysis on operational costs

To analyze the impact of different forecasting and rolling horizon methods, we run

several different scenarios where the performance is measured by the average weekly costs. For each scenario, we run 10 replications and this is sufficient for us to draw the conclusion because the standard deviation is small. All results reported are normalized. The details of scenarios are given as follows.

Scenario 1: We set planning horizon as 12 weeks and we assume that all related information is known. 12 weeks is longer than the typical planning period for operational ECR decisions which is usually 3 weeks.

Scenario 2: In this scenario, the planning horizon is set to be 3 weeks. Demand and supply of empty containers and the in-transit flow from other regions to Asia in these 3 weeks are known. To deal with the dynamic environment, the model is run in a rolling horizon manner; i.e., the model is run at the beginning of each week considering all information in the following 3 weeks. Decisions for the current week are applied and the inventory status at various ports and services are updated. The model is run every week until it reaches the 10th week.

Scenario 3: Similar to scenario 2, the length of the planning horizon is also 3 weeks, but the information of the last two weeks is based on forecasts. The forecast values are based on the average historical data.

Scenario 4: This is similar to scenario 3 except that the length of the planning horizon is reduced to 2 weeks which means only the second week forecast is used.

Scenario 5: This is similar to scenario 2 except that the length of the planning horizon is one week. Only the information in the current week is considered.

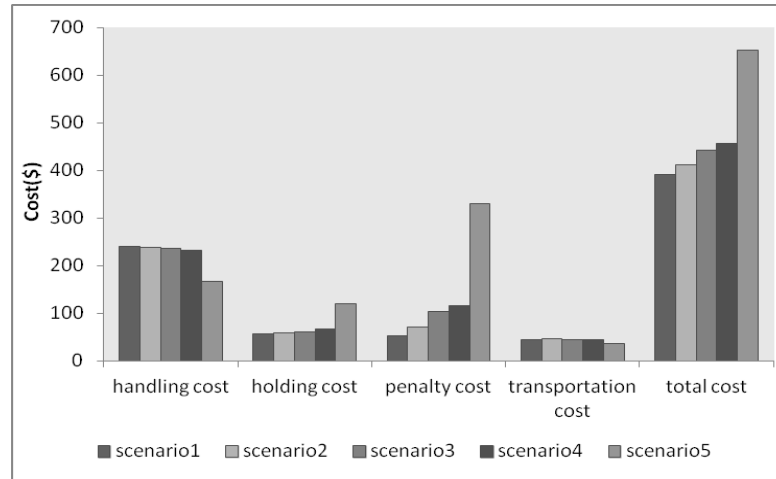


Figure 3.3 Operational costs under different lengths of planning horizon

As scenario 1 takes the whole events in the whole planning horizon into consideration, it serves as a benchmark for us to compare the performances among other scenarios. Solving this scenario will give a lower bound to other scenarios. As shown in Figure 3.3, the total cost in scenario 5 is much higher than that of other scenarios. Handling cost and transportation cost are lower in scenario 5 which means fewer empty containers are repositioned when the planning horizon is one week. However, holding cost and penalty cost in scenario 5 are much higher than those in other scenarios. It indicates that there are a lot of mismatches between demand and supply and these mismatches lead to high holding cost and penalty cost. Comparing scenario 5 with other scenarios, we find that the looking forward method with forecasting data would be helpful to reduce the mismatch.

Table 3.2 shows the operational costs of the scenario 2. The handling cost, the holding cost, the penalty cost and the transportation cost accounts for 58%, 14%, 17% and 11% of the total cost respectively. The handling cost has the largest proportion of the total cost. This result is consistent with the actual situation in the

shipping industry. It indicates that improving the efficiency of port handling activities could be the most important factor in reducing the empty container management cost. We also find that 61% of the handling cost is spent on allocating the inter-region empty containers to Asia ports while 39% of the handling cost is spent on intra-Asia ECR. This shows that empty containers repositioned from other regions are the main source of empty containers in Asia. Although Asia is the most important demand region for empty containers at the aggregated level, the intra-Asia region repositioning still plays an important role in reducing the overall costs due to the fact that the inter-region repositioning takes a longer lead time to respond to the demand.

Table 3.2 Operational costs of scenario 2

handling cost	holding cost	penalty cost	transportation cost	total cost
\$237	\$58	\$70	\$45	\$410
58%	14%	17%	11%	100%

We also compare the model with the simple rule adopted by the company. Under this simple rule, the decisions on how to satisfy the demands at each port are made sequentially. First, the shipping company starts with the inter-region services and looks at one service at a time. The empty containers at these inter-region services are used to satisfy the demands at the ports. After this allocation, any unsatisfied demand is fulfilled using the intra-Asia service, i.e., to move the containers from those ports which have extra empty containers to the ports which need empty containers. Again this is done sequentially. We have developed a modified model to capture some of

the traits of these rules. The modified model solves the problem in two phases. In phase 1, we only allocate containers that are at the inter-region services. The modified model is an optimization model which is similar to our proposed model except that the intra-Asia services are not considered. After the problem in phase 1 is solved, we update the status of the inventory and the outstanding demands at various ports. This information is used in phase 2 when we determine the empty containers that should be repositioned within ports in Asia. These decisions are also obtained using our proposed optimization model except that we do not consider the inter-region services. Note that this modified model should be better than the simple rule adopted by the company because the simple rule adopted by the company does not use optimization and it makes reposition decisions for one service at a time.

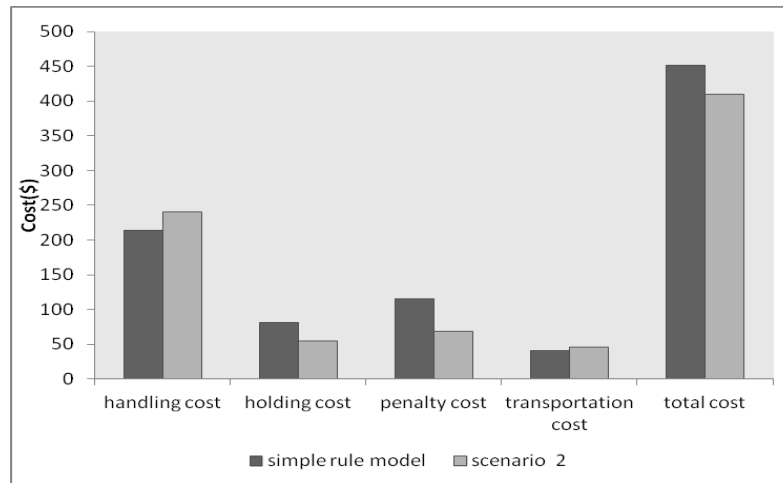


Figure 3.4 The simple rule model vs. scenario 2 of our model

Figure 3.4 compares the modified simple rule model with the scenario 2 of our model. In both models, we assume all information in the following three weeks is known. As shown in Figure 3.4, the total cost in the simple rule model is 10% higher than our proposed model. Although the handling cost for our proposed model is

slightly higher than the modified simple rule model, the penalty and the holding costs are much lower. It shows that our model may improve the ECR as it considers all the alternatives simultaneously.

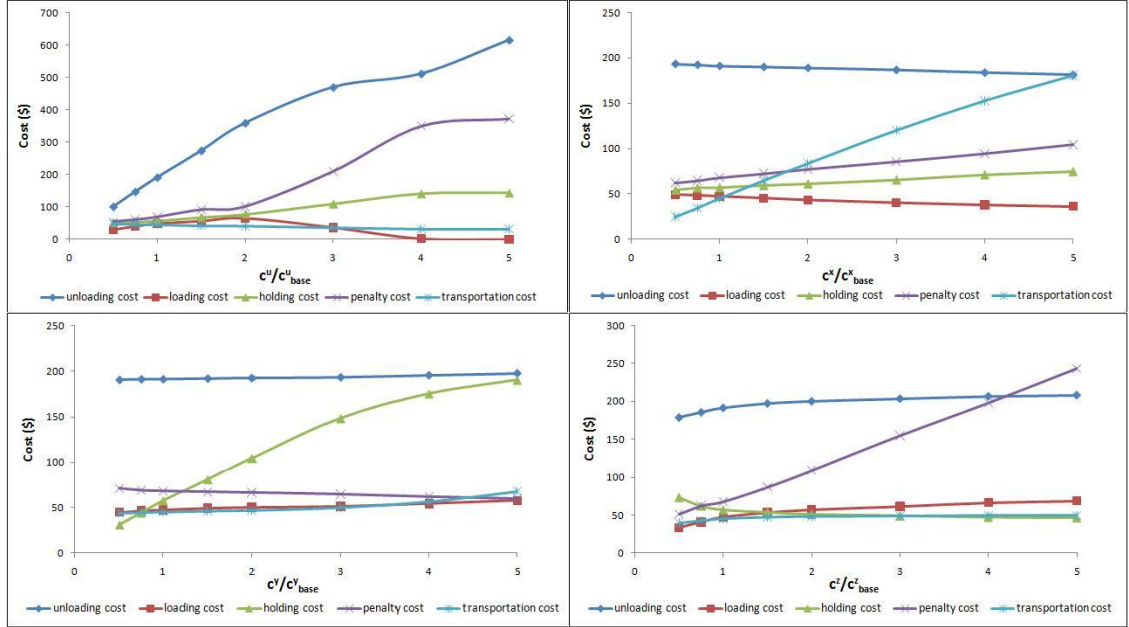


Figure 3.5 Sensitivity analysis on cost parameters

To analyze the effect of changing cost parameters on operational costs, sensitivity analysis on cost parameters is also conducted. Figure 3.5 shows how the operational costs are affected by the handling cost parameter $c_{t,i,k}^u$ and $c_{t,i,k}^w$, the holding cost parameter $c_{t,i,k}^y$, the transportation cost parameter $c_{t,s,v,k}^x$ and the penalty cost parameter $c_{t,i,k}^z$ respectively. Note that we vary the parameters values one at a time and these values are normalized with their base values. The base values are the cost parameters used in the original scenario. From the results, it is shown that the handling cost parameter $c_{t,i,k}^u$ has a great impact on operational costs. As $c_{t,i,k}^u$ increases, the total unloading cost increases dramatically. On the other hand, when c^u / c_{base}^u increases to 4, the total loading cost decreases to 0. This indicates

that due to the high handing cost parameter, we only distribute inter-region empty containers to Asia ports and we do not perform intra-region ECR. Hence, more demands in export-dominated ports are not satisfied and inventory in the import-dominated port increases. As a result, the total penalty cost and the total holding cost increase sharply.

3.3.3 Analysis on transshipment hub

In our numerical example, we observe that there is 39% of intra-Asia ECR even though many ports in this region are export-dominated ports. This implies that some ports in this region play the transshipment role to other export-dominated ports. To identify the potential transshipment hubs for ECR, we will observe the intensity of transshipment activities of empty containers at ports. It is equal to the loading activities minus the net supply of the empty container. The net supply is the remaining empty containers that returned by customers after satisfying the demand at that port. For an import-dominated port, the net supply is equal to the supply (customer returns) minus the demand of the empty container at that port. If the port is an export-dominated port, the net supply is zero.

Five Asia ports with the largest number of average transshipment activities are chosen based on the results of our model (scenario 2). The order of these five ports is Kaohsiung, Singapore, Pusan, Tokyo, and Port Klang with 815, 556, 382, 242, and 203 transshipment activities respectively. It is observed that the average

transshipment activities in Kaohsiung and Singapore are much higher than other ports, and their activities constitute to 52.2% of the total transshipment activities in Asia. By analyzing the network and outputs of our model, we summarize two significant factors that affect the transshipment activities in an export-dominated region. (1). Connectivity. All these top five ports have good connectivity in the transportation network. In this network, the number of weekly port calls in Kaohsiung, Singapore, Pusan, Tokyo, and Port Klang is 12, 37, 14, 9, and 9 respectively. Among these ports, Singapore has the highest connectivity. (2). Geographical location. For Kaohsiung port, its connectivity is much lower than Singapore, and even lower than Pusan. However, transshipment activities in Kaohsiung are much more than other ports. The main reason is that Kaohsiung is close to many export-dominated ports. By analyzing the travelling sequence of the services passing through Kaohsiung, we find that Kaohsiung is directly followed by some important export-dominated ports, i.e., Hong Kong, Yantian, Shanghai,

Table 3.3 Effects of ship capacity on transshipment activities

$\gamma/\gamma_{\text{base}}=0.5,$ $\sigma/\sigma_{\text{base}}=0.5$	Pusan 407	Kaohsiung 406	Yokahama 345	Tokyo 324	Singapore 284
$\gamma/\gamma_{\text{base}}=0.75,$ $\sigma/\sigma_{\text{base}}=0.75$	Kaohsiung 808	Singapore 704	Pusan 405	Tokyo 296	Yokahama 228
$\gamma/\gamma_{\text{base}}=1,$ $\sigma/\sigma_{\text{base}}=1$	Kaohsiung 815	Singapore 556	Pusan 382	Tokyo 242	Port Klang 203
$\gamma/\gamma_{\text{base}}=1.5,$ $\sigma/\sigma_{\text{base}}=1.5$	Kaohsiung 809	Singapore 431	Pusan 342	Tokyo 234	Port Klang 223
$\gamma/\gamma_{\text{base}}=2,$ $\sigma/\sigma_{\text{base}}=2$	Kaohsiung 834	Singapore 431	Pusan 342	Tokyo 232	Port Klang 208
$\gamma/\gamma_{\text{base}}=5,$ $\sigma/\sigma_{\text{base}}=5$	Kaohsiung 829	Singapore 426	Pusan 348	Tokyo 230	Port Klang 212

Table 3.4 Effects of cost parameter on transshipment activities

$c^u/c^u_{base}=0.5$	Kaohsiung 808	Singapore 686	Pusan 454	Port Klang 241	Tokyo 245
$c^u/c^u_{base}=1$	Kaohsiung 815	Singapore 556	Pusan 382	Tokyo 242	Port Klang 203
$c^u/c^u_{base}=1.5$	Kaohsiung 778	Singapore 547	Pusan 265	Tokyo 235	Port Klang 92
$c^u/c^u_{base}=2$	Kaohsiung 853	Singapore 482	Tokyo 224	Pusan 206	Dacao 80
$c^u/c^u_{base}=5$	Bangkok 0	Belawan 0	Bugo 0	CagayandeOro 0	CatLai 0
$c^x/c^x_{base}=0.5$	Kaohsiung 861	Singapore 550	Pusan 418	Tokyo 240	Port Klang 216
$c^x/c^x_{base}=1$	Kaohsiung 815	Singapore 556	Pusan 382	Tokyo 242	Port Klang 203
$c^x/c^x_{base}=1.5$	Kaohsiung 798	Singapore 559	Pusan 371	Tokyo 246	Port Klang 94
$c^x/c^x_{base}=2$	Kaohsiung 811	Singapore 558	Pusan 356	Tokyo 248	Port Klang 173
$c^x/c^x_{base}=5$	Kaohsiung 637	Singapore 596	Tokyo 274	Pusan 269	Port Klang 106
$c^y/c^y_{base}=0.5$	Kaohsiung 887	Singapore 558	Pusan 387	Tokyo 242	Port Klang 181
$c^y/c^y_{base}=1$	Kaohsiung 815	Singapore 556	Pusan 382	Tokyo 242	Port Klang 203
$c^y/c^y_{base}=1.5$	Kaohsiung 762	Singapore 559	Pusan 385	Tokyo 245	Port Klang 219
$c^y/c^y_{base}=2$	Kaohsiung 732	Singapore 564	Pusan 390	Tokyo 244	Port Klang 232
$c^y/c^y_{base}=5$	Singapore 677	Kaohsiung 670	Pusan 412	Port Klang 304	Tokyo 239
$c^z/c^z_{base}=0.5$	Kaohsiung 681	Singapore 568	Tokyo 232	Pusan 212	Davao 76
$c^z/c^z_{base}=1$	Kaohsiung 815	Singapore 556	Pusan 382	Tokyo 242	Port Klang 203
$c^z/c^z_{base}=1.5$	Kaohsiung 876	Singapore 579	Pusan 449	Tokyo 251	Port Klang 234
$c^z/c^z_{base}=2$	Kaohsiung 827	Singapore 664	Pusan 461	Tokyo 250	Port Klang 243
$c^z/c^z_{base}=5$	Singapore 897	Kaohsiung 676	Pusan 466	Port Klang 290	Jakarta 261

Chiwan and Laem Chabang.

We also conduct sensitivity analysis on transshipment activities in each port by changing ship residual capacity and cost parameters. In Table 3.3, we test the effects of ship capacity on transshipment activities. Five ports with the largest number of transshipment activities are presented in each scenario. In Table 3.4, we test the effects of cost parameters on transshipment activities. We find that Kaohsiung, Singapore, Pusan, Tokyo, and Port Klang have the largest number of transshipment activities in most scenarios and the ranking of these five ports is robust. Reliability to the change of ship capacity and cost parameters in these ports is very high. By comparing the four cost parameters, we also find that the handling cost parameter has the highest impact toward the transshipment activities.

3.4 Summary

This chapter presents a time space network to model the operational level ECR problem. This model aims to minimize the total empty repositioning cost. Solutions could be provided within a few seconds to fulfill the frequent need of making repositioning decisions. We also develop a decision support tool based on the model to facilitate practical application of our results in the shipping industry.

We conduct numerical experiments based on the actual services of an ocean liner which consist of 49 ports and 44 services. The results show that the handling cost has the largest contribution and is also the most sensitive parameter to the total

cost. This result indicates that the container operators should pay more attention to improve the efficiency of empty container handling at ports in order to reduce the ECR cost. Moreover, we find that the total operation cost can be greatly reduced by using the looking forward method with forecasting. Compared with the modified simple rule model, the proposed model gives as high as 10% improvement in the total cost for the scenarios that we have run. Through the numerical runs, Kaohsiung and Singapore are identified as potential transshipment hubs for this ocean liner based on their transshipment activities. Moreover, from the sensitivity analysis, we find that the positions of Kaohsiung and Singapore as transshipment hubs are also robust to the change of ship capacity and cost parameters. In our model, we adopt a rolling horizon method to handle uncertainties in the forecasting data related to the supply, the demand and the ship available capacity. An extension of this work is to develop a stochastic model by considering these uncertainties. Moreover, we can integrate the inter-region decisions with this stochastic model. Another possible direction for future research is to incorporate inland empty container transportation.

Chapter 4 A TWO-STAGE STOCHASTIC MODEL FOR EMPTY CONTAINER REPOSITIONING WITH UNCERTAINTY

In the maritime transportation, container operators have to deal with some uncertain factors like the real transportation time between two ports/deports, future demand and supply, the in-transit time of returning empty container from customers, and the available capacity in vessels for empty containers transportation, etc. In shipping industry, ocean liners usually keeps historical data on some uncertain parameters. Based on the data, distributions of these parameters in the future are able to be forecasted. Random scenarios could be generated based on these distributions.

Interview with a shipping company reveals that weekly container shipping decisions requires forecast of future demands, remaining vessels' capacities, and supply. Due to the dynamically changing environment and the low forecasting accuracy in container shipping industry, the forecasting has to be adjusted when new information is updated. Thus, we develop a two-stage stochastic model in rolling horizon policy to deal with the dynamically changing forecasting. As a widely used sampling method, Sample Average Approximation (SAA) avoids the need to approximate the value function, and it is selected to solve the stochastic Empty Container Repositioning (ECR) problem with multiple scenarios in this study.

4.1 Problem description

Table 4.1 The port rotation of service APX (westbound)

Port	City	Transit Day	Arrive	Transit Day	Depart
F96	New York	-	-	00	Sat
KFK	Norfolk	01	Sun	02	Mon
KCS	Charleston	03	Tue	04	Wed
MIT	Manhandle	07	Sat	08	Sun
OAJ	Balboa	08	Sun	09	Mon
SPQ	San Pedro	16	Mon	17	Tue
OAK	Oakland	19	Thu	20	Fri
TKY	Tokyo	30	Mon	30	Mon
KOB	Kobe	31	Tue	31	Tue
CIW	Chiwan	34	Fri	35	Sat
HKG	Hong Kong	35	Sat	36	Sun
KAO	Kaohsiung	36	Sun	37	Mon
P2H	Busan	39	Wed	-	-

Due to the global trade imbalance, containers in import-dominated regions like Europe have to be repositioned to export-dominated regions like Asia. In this case, request for containers have to be placed in advance, so containers in the import-dominated region could be loaded to vessels and then be transported to the export-dominated region. However, the lead times of inter-region orders are usually long. For example, in the port rotation of the westbound of Atlantic Pacific Express (APX) shown in Table 4.1, the transit day from New York to the first arrived port in Asia (Tokyo) is more than 30 days. The order for inter-region empty containers from North America to Asia should be placed about 4 weeks in advance, so surplus empty containers could be loaded to vessels from New York, Norfolk, Charleston, Manhandle, Balboa, San Pedro and Oakland in turn according to the order. In actual situation, whether the order could be satisfied by the import-dominated region

depends on the empty container inventory in import-dominated ports as well as the available capacity in vessels for ECR. In this study, we consider each region separately. ECR decisions related to ports in the target region are made, while the coordination with other regions is not considered. We only consider the information in the target region when decisions on the number of empty container into/out of the target region are made.

Because of the long lead time of the inter-region empty containers, ocean liners usually make ordering decisions depending on forecasting. Due to the booking system used in the maritime transportation, demand, supply and the available ship capacity in the near future could be forecasted accurately. However, it is difficult to obtain accurate forecasting for more than one or two weeks. Currently, container operators make decisions based on the nominal forecast value. Because of the difference between the expected value and the realized value, inefficient solutions may be produced.

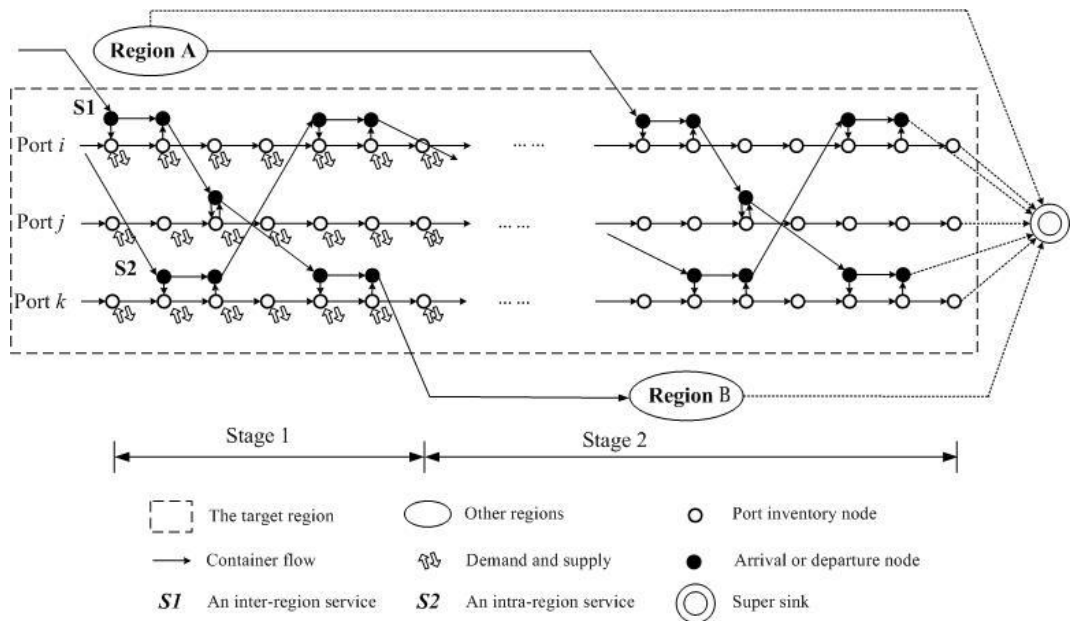


Figure 4.1 The two-stage time space network

This study aims to make short-term decisions and we also consider the influence of these decisions in the following several weeks. A time space network of the ECR problem is shown in Figure 4.1. To incorporate both deterministic information and uncertain information, we divide this network into two parts. All parameters in stage 1 are known, while some parameters in stage 2 are unknown when decisions in stage 1 are made. In Figure 4.1, *SI* is an inter-region service which travels from region A to the target region and then travels to region B. At the beginning of stage 1, container operators make ECR decisions in the target region. Besides, decisions on inter-region flows of empty container are also made.

4.2 Problem formulation

If all parameters in the planning horizon are known, the formulation of the deterministic model for ECR is similar to **P1**, which has been introduced in Chapter 3. As decisions on inter-region flows of empty container are also made, we have to consider longer planning horizon. Besides, the number of empty container into/out of the target region is a decision variable instead of a given parameter.

4.2.1 Modeling assumptions

In this study, we develop a stochastic programming model which takes into account four uncertain parameters, i.e., the demand (the empty containers that picked up by the customers to load cargos), the supply (the empty containers that returned by the

customers), the available ship space capacity, and available ship weight capacity for empty containers. Other uncertain factors like the transportation time between two ports are not considered. We also do not consider container substitution in this study. We assume that service schedule is given and fixed in the planning horizon. This assumption is valid as the planning horizon of our operation model is short (several weeks), and the service schedule is not changed frequently. Note that we do not make decisions on laden container transportation in this study. As laden container transportation problem and ECR are usually considered separately in current shipping industry, and laden container has higher priority, our model is to make ECR decisions after the laden container transportation is planned.

4.2.2 Notations

- Ω The set of all possible scenarios;
- ω A scenario that is unknown when decisions at stage 1 are made, but that is known when the decisions at stage 2 are made ($\omega \in \Omega$);
- $\xi(\omega)$ Parameters of the uncertain variables (demand, supply, residual ship weight capacity and residual ship space capacity) in scenario ω ;
- x_1 Decisions at stage 1;
- $x_2(\omega)$ Decisions for scenario ω at stage 2 given x_1 ;
- c_1, c_2 The cost vector at stage 1 and stage 2 respectively;
- A_1, B_1 The coefficient matrices of x_1 ;

A_2, B_2 The coefficient matrices of x_2 ;

a_1 The RHS of constraint (4.2);

$a_2(\omega)$ The RHS of constraint (4.6);

v The vector of ending container states of stage 1. It is the empty container inventory at each port and at each vessel at the end of stage 1 (the number of elements in v is R , $v = \{v_1, v_2, \dots, v_R\}$);

$v(x_1)$ The vector of initial container states of stage 2 given x_1 .

4.2.3 Model formulation

To incorporate uncertain parameters, i.e., demand, supply, residual ship space capacity, and residual ship weight capacity, we develop a two-stage stochastic model. This model is run in a rolling horizon manner. ECR decisions are made at the beginning of stage 1 (in this study, stage 1 is the first week) and will be made again when new information is collected. A two-stage stochastic model for ECR (defined as **P2**) is formulated as follows.

$$\mathbf{P2: Stage1:} \quad \min \quad c_1 x_1 + E_p[Q(x_1, \xi(\omega))] \quad (4.1)$$

Subject to

$$A_1 x_1 = a_1 \quad (4.2)$$

$$B_1 x_1 = v \quad (4.3)$$

$$x_1 \geq 0 \quad (4.4)$$

P2-Stage 2: For a realized scenario ω , we have

$$Q[x_1, \xi(\omega)] = \min c_2 x_2(\omega) \quad (4.5)$$

Subject to

$$A_2 x_2(\omega) = a_2(\omega) \quad (4.6)$$

$$B_2 x_2(\omega) = v(x_1) \quad (4.7)$$

$$x_2(\omega) \geq 0 \quad (4.8)$$

The objective function of stage 1 is to minimize the total operational cost in the planning horizon. $c_1 x_1$ is the operation cost at stage 1. $E_p[Q(x_1, \xi(\omega))]$ is the expected cost at stage 2. We assume that the probability distribution p on Ω is known in the stage 1. $Q(x_1, \xi(\omega))$ is the objective function of stage 2, which is the operational cost at stage 2 given x_1 and scenario ω . Constraints (4.2) and (4.6) include the typical constraints of ECR problem as shown in the deterministic model **P1**, i.e., ship capacity constraint, service flow constraint, and port flow constraint. Constraint (4.3) is to set the end container states of stage 1, which are also the initial container states of stage 2. Constraint (4.7) is to set the initial container states of stage 2. The stage 1 model and stage 2 model are connected by the container flow between these two stages. Given x_1 , the initial container states of stage 2, $v(x_1)$, should equal to the ending container states of stage 1, v . Besides the concise formulation, the explicit form of the two-stage model **P2** is also presented in Appendix A to help readers understand the two-stage model.

4.3 Methodology - Sample average approximation

Our stochastic problem is hard to solve as it is difficult to evaluate the expected cost of stage 2 for a given x_1 , i.e., $E_p[Q(x_1, \xi(\omega))]$. It requires the solutions of a large number of stage 2 optimization problems as the number of possible scenarios in the set Ω is very large in our study. Therefore, enumeration approaches may not be feasible. In this study, we apply the SAA method in addressing the large set Ω . The basic idea of SAA method is that the expected objective function of the stochastic problem is approximated by a sample average estimate derived from a random sample and the resulting SAA problem could then be solved by deterministic optimization techniques (Kleywegt et al., 2002). Several replications with different samples are run to obtain candidate solutions. This method has been applied to solve the two-stage stochastic routing problems by Verweij et al. (2003).

A sample with N scenarios $\{\omega^1, \omega^2, \dots, \omega^N\}$ is generated according to the probability distribution p . This sample is an independently and identically distributed (i.i.d.) random sample. The SAA problem (defined as **P3**) is formulated as follows.

$$\mathbf{P3:} \quad \min \quad c_1 x_1 + \frac{1}{N} \sum_{n=1}^N [c_2 x_2(\omega^n)] \quad (4.9)$$

Subject to (4.2), (4.3); and (4.6), (4.7) for $n=1, 2, \dots, N$

$$v \geq 0, x_1 \geq 0, x_2(\omega^n) \geq 0 \quad \text{for } n=1, 2, \dots, N \quad (4.10)$$

Let z^* denotes the optimal value of the true problem **P2**. The estimates of lower bound and upper bound on the optimal value z^* of the true problem **P2** could be obtained by using sampling methodology. Discussion on these estimates has been presented by previous studies of Norkin et al. (1998a) and Mak et al. (1999). Denote

the optimal value of the SAA problem (4.9) by \hat{z}_N . It is well known that $E(\hat{z}_N) \leq z^*$, where $E(\hat{z}_N)$ is the expected perceived cost. To estimate the expected value $E(\hat{z}_N)$, we can generate M independent samples equally with N scenarios. By solving the corresponding M independent SAA problems, we can get M candidate solutions $\hat{x}_N^1, \hat{x}_N^2, \dots, \hat{x}_N^M$ and the objective values $\hat{z}_N^1, \hat{z}_N^2, \dots, \hat{z}_N^M$. Then compute the value

$$\bar{z}_{N,M} = \frac{1}{M} \sum_{j=1}^M \hat{z}_N^j \quad (4.11)$$

The estimate $\bar{z}_{N,M}$ provides an unbiased estimate of $E(\hat{z}_N)$. As $E(\hat{z}_N) \leq z^*$, the estimate $\bar{z}_{N,M}$ provide a statistical lower bound of the true optimal value of the problem **P2**.

To evaluate the M candidate solutions $\hat{x}_N^1, \hat{x}_N^2, \dots, \hat{x}_N^M$, we consider generating an independent sample with N' scenarios, where N' is usually much larger than N . Let

$$\hat{z}_{N'}(\hat{x}_N^j) = c_1 \hat{x}_N^j + \frac{1}{N'} \sum_{i=1}^{N'} [Q(\hat{x}_N^j, \xi(\omega^i))] \quad (4.12)$$

where $\hat{z}_{N'}(\hat{x}_N^j)$ is the actual cost estimate corresponding to the candidate solution of \hat{x}_N^j . It is natural to take \hat{x}^* as the optimal solution of the SAA problem which provides the smallest estimated actual objective value,

$$\hat{x}^* \in \arg \min \{ \hat{z}_{N'}(\hat{x}_N^j) : \hat{x}_N \in \{ \hat{x}_N^1, \hat{x}_N^2, \dots, \hat{x}_N^M \} \} \quad (4.13)$$

As \hat{x}^* is a feasible solution of the stochastic ECR problem, $\hat{z}_{N'}(\hat{x}^*)$ gives an estimate of the upper bound of the true optimal objective value of **P2**. And the optimality gap could be estimated as

$$\hat{z}_{N'}(\hat{x}^*) - \bar{z}_{N,M} \quad (4.14)$$

4.4 Computational studies

This experimental work attempts to quantify the quality of the solution. Lower bound and upper bound of the solutions are estimated by SAA to evaluate the quality of the solutions. Note that these bounds are for a single two-stage problem and not for the two-stage policy if it were applied over multiple time periods in rolling basis. In this section, we develop a small-scale ECR transportation network. For this small-scale model, the SAA problem could be solved directly when the scenario sample is not too large. In this case, we solve the SAA problems directly by using CPLEX11.2. However, the multi-scenario SAA problem with large-scale transportation network is difficult to solve due to its scale and complexity and thus it could not be solved directly by CPLEX11.2. Advanced algorithms have to be developed. We will present several algorithms which are based on scenario decomposition to solve the larger scale case in Chapter 5.

4.4.1 The transportation network

To evaluate the performance of the SAA method, we first generate an ECR transportation network as shown in Figure 4.2. Five ports, three services, and one type of container (twenty-foot standard container) are considered. *S1* and *S2* are inter-region services which travel between region A and region B. *S3* is an

intra-region service which travels between port 2 and port 5. The planning horizon is three weeks. We define that the first week is stage 1, and the second and the third weeks are stage 2. At the beginning of stage 1, all information in stage 1 is known, while some parameters in stage 2 are unknown. These parameters are known at the beginning of stage 2. The lead time of inter-region empty containers is one week. At the beginning of stage 1, the in-transit containers in $S1$ and $S2$ which arrive at the target region A in stage 1 are known. The inter-region containers which arrive at region A at week 2 and week 3 in $S1$ and $S2$ are ordered at week 1 and week 2 respectively. The service schedules are given in Figure 4.2. Demand and supply at ports as well as ship capacity (space and weight capacity) at the beginning of stage 2 are generated randomly according to given distributions. Detailed information about random variables generation and cost parameters setting are presented in Appendix B.

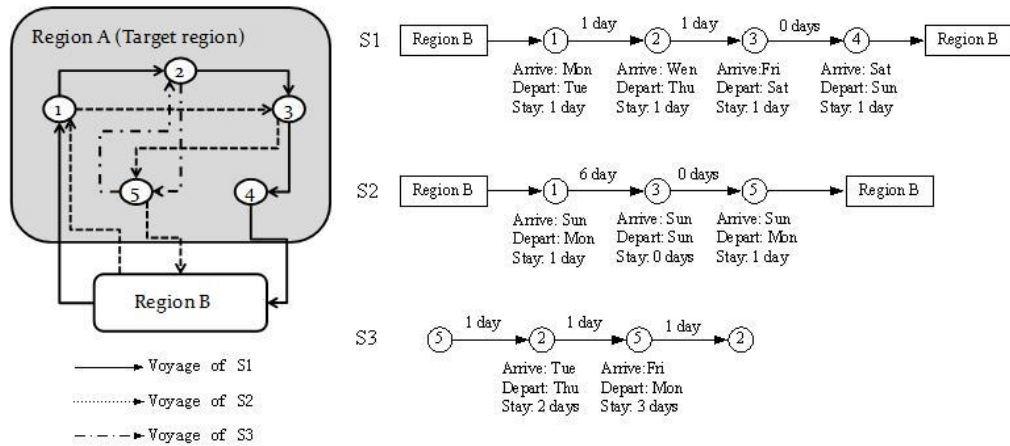


Figure 4.2 A network with three services and five ports

4.4.2 Results of the sample average approximation

We apply the SAA method to solve the two-stage stochastic problem of the small-scale case. We can solve it directly by using CPLEX11.2 when the sample size N is not too large ($N < 1000$). We set the sample size N as 100. The number of scenarios to evaluate the solution N' is set to be 1000. Replication number is set to be 20. The performance of the SAA method ($N = 100$) for the small-scale case is examined with the key results shown in Table 4.2. As shown in Table 4.2, the optimality gap $\hat{z}_{N'}(\hat{x}^*) - \bar{z}_{N,M}$ is 36.74 (1.09% of $\hat{z}_{N'}(\hat{x}^*)$) which is quite small. The small optimality gap implies that the SAA method can provide solutions with good quality.

Table 4.2 Results of the SAA method (small-scale case, $N=100$, $N'=1000$, $M=20$)

	statistical lower bound	estimated objective value	optimality gap
Estimate	$\bar{z}_{N,M}$	$\hat{z}_{N'}(\hat{x}^*)$	$\hat{z}_{N'}(\hat{x}^*) - \bar{z}_{N,M}$
Value	3326.58	3363.32	36.74(1.09%)

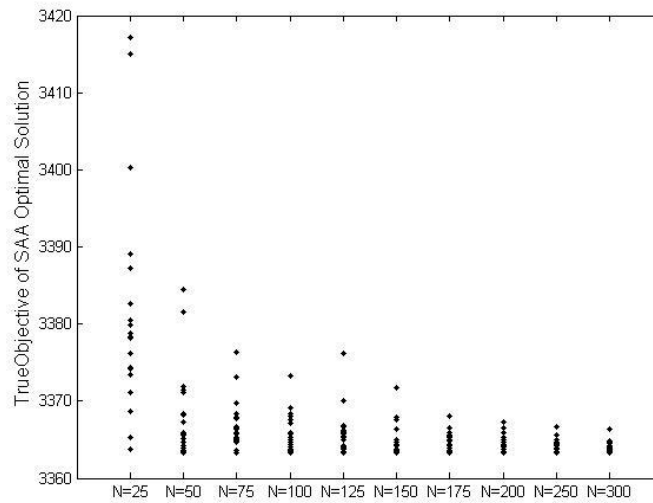


Figure 4.3 Improvement of $\hat{z}_{N'}(\hat{x}_N^j)$ with the increase of sample size N

Note that the accuracy of the approximation of the SAA method depends on the sample size N . Therefore, we analyze the performance of SAA method by changing the sample size N . Figure 4.3 shows the improvement of objective value $\hat{z}_{N'}(\hat{x}_N^j)$ of SAA optimal solutions \hat{x}_N^j with the increase of the sample size N . We consider sample size $N=20, 50, 75, 100, 125, 150, 175, 200, 250$, and 300 . The number of replications of each sample size is 20. As shown in Figure 4.3, good solutions could be obtained when sample size is 20. However, the variance of the estimate of the objective value is large when the sample size is small. In order to obtain good and reliable solutions, the results suggest that it is better to increase the sample size until the variance of the estimate of the objective value is sufficiently small.

4.4.3 Deterministic model vs. stochastic model

To evaluate the reduction in operational cost by considering uncertainties in the ECR problem, the solutions of a stochastic model are compared with the solutions of a deterministic model. For the deterministic model, we generate a scenario in which all parameters in stage 1 are the same with the parameters in stage 1 of the stochastic model. The parameters in stage 2 of this scenario are set to be the means of these parameters. By solving the deterministic model, we can get a solution \bar{x}_1 . We then could evaluate this solution by N' scenarios. The estimated objective value with the solution from the deterministic model is denoted by $\hat{z}_{N'}(\bar{x})$, where

$$\hat{z}_{N'}(\bar{x}_1) = \min \quad c_1 \bar{x}_1 + \frac{1}{N'} \sum_{n=1}^{N'} Q[\bar{x}_1, \xi(\omega^n)] \quad (4.15)$$

The estimated objective value provided by a deterministic model $\hat{z}_N(\bar{x}_1)$ and the estimated objective value provided by a stochastic model are shown in Table 4.3. For the small-scale case, the estimated objective value provided by the stochastic model is 3.4% lower than the estimated objective value provided by the deterministic model. This result shows that our stochastic model can reduce the ECR operation cost compared to the deterministic model.

Table 4.3 Deterministic model vs. stochastic model (small-scale case, single period)

	cost of deterministic model(\$)	cost of stochastic model (\$)	improvement
small-scale case	3481.04	3363.32	3.4%

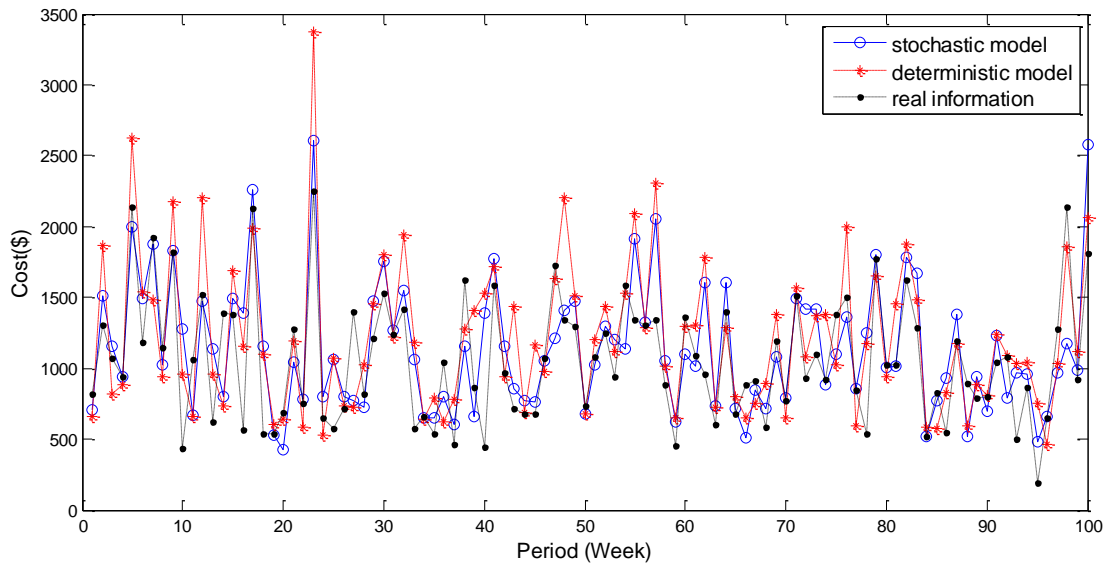


Figure 4.4 Weekly cost of the stochastic model and deterministic model overtime

The performance of the two-stage stochastic model overtime is also studied in Figure 4.4. For the small-scale network, we run the SAA problem in rolling horizon basis for 100 periods. And we compare the costs of the stochastic model and the

deterministic model. We also consider a scenario with all the real information in the planning horizon (three weeks) is known and let this scenario serve as a benchmark. We find that the average weekly operation cost of the stochastic model is 1138.96, which is 6.2% lower than the average weekly operational cost of the deterministic model. It indicates that this policy could reduce the cost in rolling basis.

4.5 Summary

In order to incorporate uncertainties in ECR problem, we build a two-stage stochastic model with uncertain demand, supply, residual ship space capacity, and residual ship weight capacity. We apply the SAA method to solve this stochastic problem. To estimate the performance of the SAA method, we solve a small-scale case. The small optimality gap indicates that the SAA method could provide good solution for the two-stage stochastic ECR problem. By comparing solutions from the deterministic model and the stochastic model, it is implied that the operational cost for ECR could be reduced by considering uncertainties. However, the multi-scenario SAA problem with large-scale transportation network is difficult to solve due to its scale and complexity and thus it could not be solved directly by CPLEX11.2. Advanced algorithms have to be developed. We will present several algorithms which are based on scenario decomposition to solve the larger scale case in Chapter 5.

In this study, we do not consider information in other regions when we make

ECR decisions in the target region. In the future, we may relax this assumption and consider the coordination amongst several regions. Besides, as the two-stage model lives with the assumption that we know the future at the end of the first stage, it is unable to handle sequential decisions. We may extend the two-stage problem to multi-stage problem to solve the maritime ECR problem in our future study.

Chapter 5 PROGRESSIVE HEDGING STRATEGY FOR STOCHASTIC EMPTY COTAINER REPOSITIONING

The multi-scenario Sample Average Approximation (SAA) problem with large-scale transportation network is difficult to solve due to its scale and complexity and thus it could not be solved directly by CPLEX11.2. To successfully apply the SAA method to solve the stochastic Empty Container Repositioning (ECR) problem, efficient algorithms should be developed to solve the approximated stochastic ECR problem, through the use of scenarios. In this chapter, we consider applying scenario aggregation by combining the approximate solution of the individual scenario problem.

In Section 5.2, we present two Progressive Hedging Approximation (PHA)-based algorithms to solve the multi-scenario SAA problem with given sample size. The progressive hedging strategy is suitable for the multi-scenario SAA problem by providing an effective decomposition. It is widely applied to solve stochastic programming problems like fisheries management problem, stochastic inventory routing problem and stochastic network design problem. In Section 5.3, we will extend to PHA-based algorithms with sequential sampling, i.e., new scenarios are generated and added to current sample iteratively. In this case, the sample size may not have to be determined in advance.

5.1 Scenario decomposition

In order to separate scenarios, we can model the original SAA problem **P3** from the scenario-focused perspective (defined as **P4**),

$$\mathbf{P4:} \quad \min \quad \frac{1}{N} \sum_{n=1}^N [c_1 x_1(\omega^n) + c_2 x_2(\omega^n)] \quad (5.1)$$

Subject to

$$A_1 x_1(\omega^n) = a_1 \quad \text{for } n=1, 2, \dots, N \quad (5.2)$$

$$B_1 x_1(\omega^n) = v(\omega^n) \quad \text{for } n=1, 2, \dots, N \quad (5.3)$$

$$A_2 x_2(\omega^n) = a_2(\omega^n) \quad \text{for } n=1, 2, \dots, N \quad (5.4)$$

$$B_2 x_2(\omega^n) = v(\omega^n) \quad \text{for } n=1, 2, \dots, N \quad (5.5)$$

$$v(\omega^n) = \bar{v} \quad \text{for } n=1, 2, \dots, N \quad (5.6)$$

$$\bar{v} \geq 0, \quad v(\omega^n) \geq 0, \quad x_1(\omega^n) \geq 0, \quad x_2(\omega^n) \geq 0 \quad \text{for } n=1, 2, \dots, N \quad (5.7)$$

P4 is a very large-scale linear programming model. Constraints (5.6) are to ensure that the ending states in stage 1 of different scenarios to be the same. These constraints are added to produce *implementable* solutions in stage 1. Augmented Lagrangian is applied to relax constraint (5.6). After applying the Augmented Lagrangian, we have

P5:

$$\min \frac{1}{N} \sum_{n=1}^N [c_1 x_1(\omega^n) + c_2 x_2(\omega^n) + \sum_{r=1}^R \lambda_r(\omega^n) (v_r(\omega^n) - \bar{v}_r) + \frac{1}{2} \rho \sum_{r=1}^R (v_r(\omega^n) - \bar{v}_r)^2] \quad (5.8)$$

Subject to

$$(5.2) - (5.5) \text{ and } (5.7)$$

where R is the number of elements in \bar{v} , and $\lambda_r(\omega^n)$ ($r=1, 2, \dots, R$ and $n=1, 2, \dots, N$)

are the Lagrangian multipliers.. $\frac{1}{2} \rho \sum_{r=1}^R (v_r(\omega^n) - \bar{v}_r)^2$ is the cost to pay for the differences between the scenario solutions and the overall solution \bar{v} . The relaxed problem **P5** is not scenario separable. However, if the overall solution \bar{v} is given and fixed, the relaxed problem can be separated according to individual scenario. For a scenario ω , the corresponding sub-problem (defined as **P6**) given \bar{v} could be expressed as:

$$\mathbf{P6:} \quad \min[c_1 x_1(\omega) + c_2 x_2(\omega) + \sum_{r=1}^R \lambda_r(\omega)(v_r(\omega) - \bar{v}_r) + \frac{1}{2} \rho \sum_{r=1}^R (v_r(\omega) - \bar{v}_r)^2] \quad (5.9)$$

Subject to

$$A_1 x_1(\omega) = a_1 \quad (5.10)$$

$$B_1 x_1(\omega) = v(\omega) \quad (5.11)$$

$$A_2 x_2(\omega) = a_2(\omega) \quad (5.12)$$

$$B_2 x_2(\omega) = v(\omega) \quad (5.13)$$

$$v(\omega) \geq 0, \quad x_1(\omega) \geq 0, \quad x_2(\omega) \geq 0 \quad (5.14)$$

The objective value of (5.9) is denoted by $g_{P6}(\omega)$. By solving the sub-problem **P6**, solutions for different scenarios may be different. In order to obtain consensus amongst these sub-problems, we apply heuristic algorithms to find an *implementable* solution for the SAA problem.

5.2 Progressive hedging approximation -based algorithms for sample average approximation problem

In order to obtain consensus amongst these sub-problems, we develop heuristic algorithms to find an *implementable* solution for the SAA problem.

5.2.1 Progressive hedging approximation -based algorithm 1

To obtain an overall solution \bar{v} , we consider the local information provided by the sub-problem of each scenario realization. Let i denotes the index of the number of iteration. By sequentially solving the sub-problem **P6** for $\omega^1, \omega^2, \dots, \omega^N$, we can get $v^i(\omega^1), v^i(\omega^2), \dots, v^i(\omega^N)$. Let

$$\bar{v}_r^i = \frac{1}{N} \sum_{n=1}^N v_r^i(\omega^n) \quad \text{for } r=1, 2, \dots, R \quad (5.15)$$

$\bar{v}^i = \{\bar{v}_1^i, \bar{v}_2^i, \dots, \bar{v}_R^i\}$ is used to identify trends and essential features among scenario solutions $v^i(\omega^1), v^i(\omega^2), \dots, v^i(\omega^N)$. As a reference point, (5.15) provides an overall solution in the i^{th} iteration. However, this reference point may not be feasible for the original SAA problem. To produce a feasible *implementable* solution in the i^{th} iteration (denoted as $v^{i, \text{feasible}}$), we can select one solution among $v^i(\omega^1), v^i(\omega^2), \dots, v^i(\omega^N)$. In this study, we select the solution with the maximum objective value of **P6**, i.e.,

$$v^{i, \text{feasible}} \in \arg \max \{g_{P6}^i(v^i(\omega^n)) : v^i \in \{v^i(\omega^1), v^i(\omega^2), \dots, v^i(\omega^N)\}\} \quad (5.16)$$

As $v^{i, \text{feasible}}$ is a feasible solution, the objective value of **P4** with constraints $v(\omega^n) = v^{i, \text{feasible}}, \text{ for } n=1, 2, \dots, N$ provides an upper bound for the objective value

of the SAA problem **P4**. If \bar{v} is given, the SAA problem **P4** could be separated into the following sub-problem (defined as **P7**). For scenario ω ,

$$\mathbf{P7:} \quad \min \quad c_1 x_1(\omega) + c_2 x_2(\omega) \quad (5.17)$$

Subject to

$$(5.10) - (5.14), \text{ and } v(\omega) = v^{i, \text{feasible}}$$

The objective value of (5.17) is denoted by $g_{P7}(\omega)$. And we define the objective value of **P7** with $v(\omega) = v^{i, \text{feasible}}$ as $g_{P7}^i(v^{i, \text{feasible}}, \omega)$. So one upper bound of the SAA problem **P4** in the i^{th} iteration is denoted as:

$$\text{UB}^i = \frac{1}{N} \sum_{n=1}^N g_{P7}^i(v^{i, \text{feasible}}, \omega^n) \quad (5.18)$$

If the feasible solution $v^{i, \text{feasible}}$ is calculated iteratively and we keep the best upper bound, we always have a solution that is best until now at hand. This proceeding in generating improving solutions is the principle of progressive hedging in optimization under uncertainties (Rockafeller and Wets, 1991).

In order to obtain consensus amongst scenario solutions, $\lambda_r(\omega^n)$ and ρ are updated iteratively. By adjusting the penalty parameters for the differences between the scenario solutions obtained and the reference point generated, the scenario solutions are then forced to converge to an overall solution. In this study, we apply the typical approach to update $\lambda_r(\omega^n)$ and ρ ,

$$\lambda_r^i(\omega^n) \leftarrow \lambda_r^{i-1}(\omega^n) + \rho^{i-1}(v_r^i(\omega^n) - \bar{v}_r^{i-1}) \quad (5.19)$$

$$\rho^i \leftarrow \alpha \rho^{i-1} \quad (5.20)$$

where $\alpha > 1$ is a given constant.

The iteration proceeds until the stopping criterion are met. In this study, we

consider two conditions. First, the differences between the scenario solutions obtained and the reference point generated are sufficient small, i.e., $\sum_{n=0}^N \sum_{r=1}^R (v_r^i(\omega^n) - \bar{v}_r^i)^2 \leq \eta$, where η is the error tolerance which should be a positive real number. The second condition is that there are L consecutive non-improving iterations. The iterative process stops if both of these conditions are met. The PHA-based algorithm 1 is summarized as follows.

The PHA-based algorithm 1

Step 1: Initialization. Set $\lambda_r(\omega^n) \leftarrow 0$ for $r=1, 2, \dots, R$ and $n=1, 2, \dots, N$. Set $\rho^i \leftarrow \rho^0$.

Step 2: Solve **P6** for each scenario and obtain $v^i(\omega^1), v^i(\omega^2), \dots, v^i(\omega^N)$.

Step 3: Calculate the reference point \bar{v}_r^i for $r=1, 2, \dots, R$.

Step 4: Calculate and evaluate $v^{i, feasible}$. Then get UB^i and update the best solution if $v^{i, feasible}$ gives the current best objective value.

Step 5: Check the stopping criterion. If the stopping criterion is satisfied, stop.

Otherwise, go to Step 6.

Step 6: Update $\lambda_r(\omega^n)$ and ρ according to (5.19), (5.20). Go back to Step 2.

5.2.2 Progressive hedging approximation -based algorithm 2

As there is a nonlinear term $\frac{1}{2} \rho \sum_{r=1}^R (v_r(\omega) - \bar{v}_r)^2$ in the objective function (5.9), we consider another way to force $v(\omega)$ to converge by augmenting the objective function. For a scenario ω , the objective function of the corresponding sub-problem

could be expressed as:

$$\min \quad c_1 x_1(\omega) + c_2 x_2(\omega) + \sum_{r=1}^R \lambda_r(\omega^n) |v_r(\omega^n) - \bar{v}_r| \quad (5.21)$$

Then, the sub-problem could be converted to linear programming easily.

$$\mathbf{P8:} \quad \min \quad c_1 x_1(\omega) + c_2 x_2(\omega) + \sum_{r=1}^R \lambda_r(\omega^n)(o_r + o'_r) \quad (5.22)$$

Subject to

$$(5.10) - (5.14)$$

$$o_r - o'_r = v_r(\omega^n) - \bar{v}_r \quad \text{for } r=1, 2, \dots, R \quad (5.23)$$

$$o_r \geq 0, o'_r \geq 0, \text{ for } r=1, 2, \dots, R \quad (5.24)$$

The multiplier $\lambda_r^i(\omega^n)$ and ρ are updated as follow:

$$\lambda_r^i(\omega^n) = \lambda_r^{i-1}(\omega^n) + \rho^i |v_r^i(\omega^n) - \bar{v}_r^{i-1}| \quad (5.25)$$

$$\rho^i = \frac{\alpha(UB^i - LB^i)}{G^i} \quad (5.26)$$

where α is a user-defined scalar, $LB^i(UB^i)$ is the lower bound (the upper bound) obtained in the i^{th} iteration. As the solutions for different scenarios may be different for the coordinated sub-problem **P8**, one lower bound which could be obtained in iteration i is denoted as

$$LB^i = \frac{1}{N} \sum_{n=1}^N g_{\text{P8}}^i(\omega^n) \quad (5.27)$$

The upper bound in the i^{th} iteration could be obtained according to (5.18). In (5.20), as $\alpha > 1$ is a given constant, the penalty ratio ρ^i increases with the number of iterations. Accordingly, sub-problem solutions are gradually forced to converge to an *implementable* solution. In (5.26), the penalty ratio is updated dynamically. If the

optimality gap between the lower bound LB^i and the upper bound UB^i is large, it indicates the current feasible solution is far from the optimal solution. In order to force it to converge to the optimal solution, we increase the penalty ratio. Otherwise, we reduce the penalty ratio. G^i is defined to estimate the differences between the scenario solutions obtained and the reference point generated in iteration i , where

$$G^i = \frac{1}{N} \sum_{n=1}^N \sum_{r=1}^R (v_r^i(\omega^n) - \bar{v}_r^i)^2 \quad (5.28)$$

As the decrease of the differences between the scenario solutions and the reference point, the penalty ratio is increased. Thus, the sub-problem **P8** is forced to converge further.

5.2.3 Computational studies

In this section, we present two case studies in different scales. For the small-scale case which has been introduced in Section 4.4, the SAA problem could be solved directly by using CPLEX11.2 when the scenario sample is not too large. In this section, we apply two progressive hedging based algorithms to solve the SAA problems of the small-scale case. The results obtained from these two approximation algorithms are compared with the results by solving the SAA problems using CPLEX11.2. Next, we consider a real-scale ECR transportation network with a large number of ports and services. The SAA problem of this large-scale case is difficult to solve directly. Thus, we apply the two approximation algorithms based on the progressive hedging strategy to solve it. We finally compare the results of the

deterministic model and stochastic model.

5.2.3.1 Case study 1: a small-scale case

The transportation network and parameter generation have been introduced in Section 4.4.1 and Appendix B. To evaluate the performance of the progressive hedging based algorithms, we compare the estimated objective values of the SAA problems provided by these two algorithms with the objective values of the SAA

Table 5.1 The performance of the progressive hedging based algorithms for small-scale cases ($L=20, \eta=2$)

Replication (M)	\hat{z}_N^j	algorithm 1 ($\rho^0=0.5, \alpha=1.04$)		algorithm 2 ($\alpha=0.05$)	
		estimated \hat{z}_N^j	$ e_m $	estimated \hat{z}_N^j	$ e_m $
1	3232.56	3236.21	3.65	3232.81	0.25
2	3152.04	3154.88	2.84	3152.23	0.19
3	3257.00	3259.20	2.20	3257.28	0.28
4	3375.73	3377.79	2.06	3376.38	0.65
5	3377.38	3380.4	3.02	3377.74	0.36
6	3335.71	3339.28	3.57	3336.27	0.56
7	3475.33	3477.67	2.34	3475.5	0.17
8	3274.44	3276.37	1.93	3274.66	0.22
9	3348.56	3351.06	2.50	3349.43	0.87
10	3305.2	3307.66	2.46	3305.65	0.45
11	3296.68	3298.03	1.35	3296.73	0.05
12	3287.15	3289.27	2.12	3287.44	0.29
13	3435.56	3438.02	2.46	3435.92	0.36
14	3337.01	3340.39	3.38	3337.22	0.21
15	3326.44	3327.61	1.17	3326.8	0.36
16	3369.03	3371.34	2.31	3369.16	0.13
17	3219.51	3222.43	2.92	3220.04	0.53
18	3385.98	3387.49	1.51	3386.82	0.84
19	3442.62	3444.68	2.06	3442.79	0.17
20	3297.71	3300.73	3.02	3297.96	0.25

problems that are solved directly by using CPLEX11.2. Table 5.1 shows the performance of the progressive hedging based algorithms ($L=20, \eta=2$). In Table 5.1, the second column is the objective value of the 20 replications of SAA problems **P3**

solved directly by using CPLEX11.2. The estimated \hat{z}_N^j provided by the two progressive hedging algorithms are also shown. By comparing \hat{z}_N^j with the estimated \hat{z}_N^j , we find that the estimate error is small. The average percentage of differences of algorithm 1 is 0.07% and the average percentage of differences of algorithm 2 is 0.01%. The result validates that these two algorithms can be applied to solve the SAA problem and can provide good approximation of the SAA problem. We also find that the algorithm 2 could provide better approximation.

5.2.3.2 Case study 2: a large-scale case

We now apply the SAA method to solve a large-scale case. We consider a real-scale network with 49 ports and 44 services which has been analyzed in Section 3.3. Some parameters about the scale of the network are shown in Table 5.2. The planning horizon is 6 weeks, where the first week is stage 1 and the 2nd week to the 6th week is stage 2. We set the half of the cycle travelling time of services as the lead time of inter-region empty containers. The cycle travelling time of these 21 inter-region services are shown under Table 5.2. Demand, supply, and ship capacity are generated based on normal distribution.

Table 5.2 Network parameters for the large-scale problem

Number of ports	49
Total services	44
Inter-region services	21
Intra-region services	23
Planning horizon	6 weeks
Types of container	4 (20 ft standard, 40 ft standard, 40 ft high cube and 45 ft standard)

****Inter-region services with cycle travelling time:***

North America-Asia services: APX(10 weeks), ESX(6 week), JAS(6 week), NYX(10 week), PCE(6 week), PCX(6 week), PNW(6 week), PS1(8 week), PS2(6 week), PSX(8 week), SAX(8 week);

Europe-Asia services: AEX(10 week), CFX(10 week), ESX(10 week), EU2(10

week), EU3(10 week), JEX(10 week), MED(10 week), MEX(10 week), NCE(10 week), SCX(10 week).

* The cycle travelling time of the inter-region services are listed in the parentheses

For this large-scale network, the SAA problems with multiple scenarios are difficult to solve directly. Therefore, we apply the progressive hedging based algorithms to solve this large-scale case. The estimated \hat{z}_N^j and estimated $\hat{z}_{N'}(\hat{x}_N^j)$ by these two progressive hedging based algorithms are shown in Table 5.3. The sample size N is set to be 30. Replication number M is set be 10. The number of

Table 5.3 The performance of the progressive hedging based algorithms for large-scale cases ($L=10$, $\eta=5000$)

Replication (M)	algorithm 1($\rho^0=0.5$, $\alpha=1.04$)		algorithm 2($\alpha=0.05$)	
	estimated	estimated	estimated	estimated
	\hat{z}_N^j	$\hat{z}_{N'}(\hat{x}_N^j)$	\hat{z}_N^j	$\hat{z}_{N'}(\hat{x}_N^j)$
1	2.210×10^7	2.221×10^7	2.209×10^7	2.220×10^7
2	2.212×10^7	2.220×10^7	2.210×10^7	2.219×10^7
3	2.213×10^7	2.221×10^7	2.212×10^7	2.220×10^7
4	2.214×10^7	2.220×10^7	2.213×10^7	2.219×10^7
5	2.255×10^7	2.221×10^7	2.253×10^7	2.220×10^7
6	2.197×10^7	2.220×10^7	2.196×10^7	2.219×10^7
7	2.219×10^7	2.220×10^7	2.217×10^7	2.219×10^7
8	2.209×10^7	2.220×10^7	2.208×10^7	2.219×10^7
9	2.198×10^7	2.220×10^7	2.197×10^7	2.219×10^7
10	2.198×10^7	2.221×10^7	2.197×10^7	2.220×10^7

Table 5.4 Results of the SAA method (large-scale case, $N=30$, $N'=300$, $M=10$)

Estimate	statistical lower bound	estimated objective value	optimality gap
	$\bar{z}_{N,M}$	$\hat{z}_{N'}(\hat{x}^*)$	$\hat{z}_{N'}(\hat{x}^*) - \bar{z}_{N,M}$
Algorithm 1	2.212×10^7	2.220×10^7	75130(0.34%)
Algorithm 2	2.211×10^7	2.219×10^7	74860(0.34%)

scenarios to evaluate the solution N' is set to be 300. We set η to be 5000 and stop to

run more iteration when the improvement of 10 consecutive iterations in objective value is less than 500 (0.002% of the estimated $\hat{z}_{N'}(\hat{x}^*)$). As shown in Table 5.3, the estimated objective value $\hat{z}_{N'}(\hat{x}_N^j)$ provided by the algorithm 2 is smaller than the estimated objective value $\hat{z}_{N'}(\hat{x}_N^j)$ provided by the algorithm 1. The results suggest that the algorithm 2 may provide better solution than the algorithm 1. One main reason is that we dynamically update the penalty ratio in the algorithm 2. Based on the estimated \hat{z}_N^j and estimated $\hat{z}_{N'}(\hat{x}_N^j)$ in Table 5.3, the results of SAA method are summarized in Table 5.4. As shown in Table 5.4, the optimal gaps $\hat{z}_{N'}(\hat{x}^*) - \bar{z}_{N,M}$ are quiet small. It indicates that we can get good solution by applying the SAA method to solve the large-scale case.

Table 5.5 Deterministic model vs. stochastic model (large-scale case, single period)

	cost of deterministic model(\$)	cost of stochastic model (\$)	improvement
large-scale case	2.245×10^7	2.219×10^7	1.2%

For the small-scale case, the estimated objective value provided by the deterministic model and the estimated objective value provided by the stochastic model has been compared in Section 4.4.3. In this section, we also compare the operational costs of the deterministic model and the stochastic model for the large-scale case. As shown in Table 5.5, for the large-scale case, the estimated objective value provided by the stochastic model is 1.2% lower than the estimated objective value provided by the deterministic model. This result shows that our stochastic model can reduce the ECR operation cost compared to the deterministic

model. Note that the deterministic model and stochastic model lead to different initial container states of stage 2. The initial container states of stage 2 affect the total operational cost of stage 2. Therefore, the longer of the stage 2, the relative improvement of the stochastic model is smaller.

5.3 Progressive hedging approximation -based algorithm with sequential sampling

In Section 5.2, we present two PHA-based algorithms to solve the SAA problem, where the sample size N of the SAA problem is given. As we know, identifying the appropriate sample size is a crucial problem in the application of the SAA method. In this section, we consider developing PHA-based algorithm with sequential sampling, i.e., we start with a small sample. Then, new scenarios are generated and added to the current sample iteratively. According to the progressive hedging strategy, the reference point, feasible solution and best solution so far are updated iteratively. The advantage of this approach is that we do not have to determine the sample size in advance. Instead, the solution could be found by continually adding new scenarios until it convergences to an optimal/sub-optimal solution. In Section 5.3.1, detail explanation of the PHA-based algorithm with sequential sampling is presented. Computational studies to evaluate the performance of the PHA-based algorithm with sequential sampling are presented in Section 5.3.2.

5.3.1 Sequential sampling

Identifying the appropriate sample size is a crucial problem in the application of the SAA method. To determine the sample size N , the trade-off between the quality of an optimal solution of the SAA problem and the computational effort should be considered. According to the computational results in Section 4.4.2, the quality of the optimal solution improves along with the increase of sample size N . In order to obtain good and reliable solutions, the results in Figure 4.3 suggest that it is better to increase the sample size until the variance of the estimate of the objective value is sufficient small. In practice, the sample size may be adjusted dynamically, depending on the results of preliminary computations.

In this section, we propose a new algorithm which is developed based on the PHA-based algorithms introduced in Section 5.2. Sequential sampling is considered in the proposed algorithm. In this case, we do not have to determine the sample size in advance. We start from an i.i.d. sample with $N1$ scenarios. At the i^{th} iteration, the total number of scenarios in consideration is denoted by N_i . According to the PHA-based algorithms introduced in Section 5.2, scenario decomposition is applied and the sub-problem for each scenario is solved according to **P6** (or **P8**). By combining the information provided by these sub-problems of the current sample, a preliminary reference point \bar{v}_k^{i,N_i} is obtained.

$$\bar{v}_k^{i,N_i} = \frac{1}{N_i} \sum_{n=1}^{N_i} v_k^i(\omega^n) \quad (5.29)$$

Next, a new sample with $N2$ scenarios is generated. Similarly, the $N2$ scenarios are

solved separately and then a reference point $\bar{v}_k^{i,N2}$ which combines the information of the $N2$ scenarios is obtained.

$$\bar{v}_k^{i,N2} = \frac{1}{N2} \sum_{n=1}^{N2} v_k^i(\omega^n) \quad (5.30)$$

The $N2$ scenarios are added to the current sample, and N_i is updated.

$$N_{i+1} = N_i + N2 \quad (5.31)$$

The overall reference point \bar{v}_k^i could be estimated by

$$\bar{v}_k^i \leftarrow w_1 \bar{v}_k^{i,N_i} + w_2 \bar{v}_k^{i,N2} \quad (5.32)$$

where w_1 and w_2 are the weight parameters of \bar{v}_k^{i,N_i} and $\bar{v}_k^{i,N2}$ respectively. As the reference point \bar{v}_k^{i,N_i} combines the information of more scenarios than $\bar{v}_k^{i,N2}$ ($N_i > N2$), and it has been updated for i iterations. The weight parameter w_1 should be larger than the weight parameter w_2 . In this study, the reference point \bar{v}_k^i is updated as

$$\bar{v}_k^i = \frac{\bar{v}_k^{i,N_i} \cdot N_i + \phi \cdot \bar{v}_k^{i,N2} \cdot N2}{N_i + \phi \cdot N2} \quad (5.33)$$

where $\phi < 1$ is a given constant. The feasible solution could be obtained according to (5.16) and the best solution until the i^{th} iteration is updated. Finally, based on the overall reference point \bar{v}_k^i , the multipliers of the sub-problem are also updated. These procedures are conducted iteratively until the stopping criterions are met. The general procedure of the proposed algorithm is summarized in Figure 5.1.

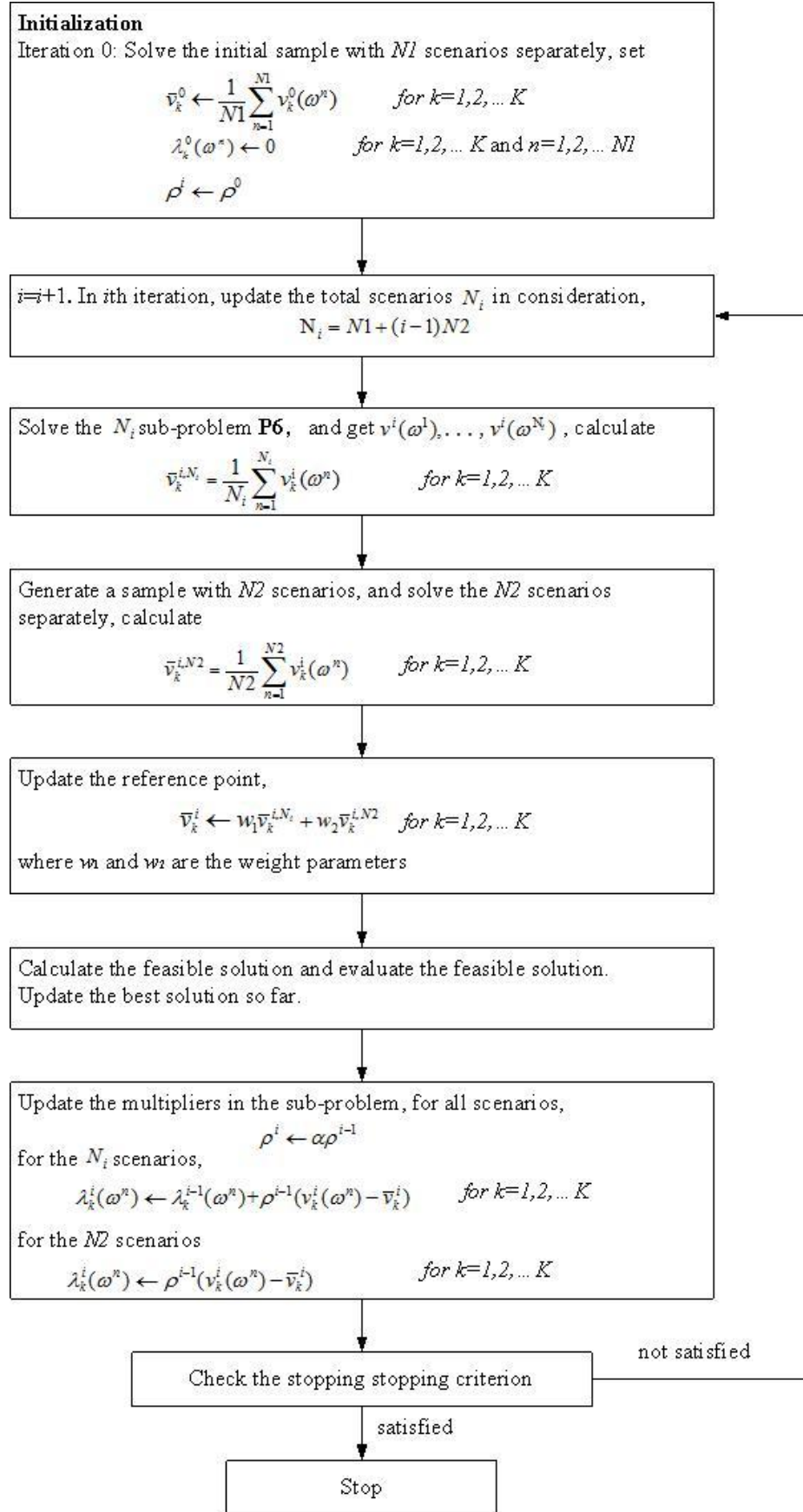


Figure 5.1 PHA-based algorithm with sequential sampling

5.3.2 Computational studies

Computational experiments are designed to evaluate the performance of the PHA-based algorithm with sequential sampling. The transportation network and parameters generation have been introduced in Section 4.4.1 and Appendix B. The number of scenarios used to evaluate the actual cost of a candidate solution N' is set to be 1000. The sub-problem formulation and the approaches to update the multipliers of the sub-problem are the same with PHA-based algorithm 1, which has been introduced in Section 5.2.1. The overall reference point is updated according to (5.33), where ϕ is set to be 0.5. Three cases are studied in this section. The replication number M of each case is 10.

Case 1: The initial sample size $N1$ is 5. The sample size of sequential sampling $N2$ is set to be 5. The number of iterations is 50.

Case 2: The initial sample size $N1$ is 5. The sample size of sequential sampling $N2$ is set to be 2. The number of iterations is 100.

Case 3: The initial sample size $N1$ is 5. The sample size of sequential sampling $N2$ is set to be 1. The number of iterations is 200.

The performance of the the PHA-based algorithm with sequential sampling for these three cases are shown in Figure 5.2 - Figure 5.7. One replication of case 1 solved by the PHA-based algorithm with sequential sampling is shown in Figure 5.2, where the estimated actual cost of the i^{th} iteration and the best objective value by the i^{th} iteration are presented. According to the results of the SAA method with fixed sample size ($N=100$) in Table 4.2, the estimated optimal objective value $\hat{z}_{N'}(\hat{x}^*)$ is

3363.32. As shown in Figure 5.2, both the estimated actual cost and the best objective value converge to the optimal value. Similarly, Figure 5.4 and Figure 5.6 show that the estimated actual cost and the best objective value of case 2 and case 3 also converge to the optimal value. It is found that the estimated actual cost in case 3 is more unstable than that in case 1 and case 2. This result is attributed to the smaller $N2$ in case 3, which makes the reference point $v_k^{i,N2}$ fluctuate. The estimated actual costs of 10 replications of case 1, case 2, and case 3 are shown in Figure 5.3, Figure 5.5, and Figure 5.7 respectively. It is shown that the solution of the proposed algorithm with sequential sampling converges to the optimal solution in all these replications. It provides clear evidence that the PHA-based algorithm with sequential sampling could provide optimal/sub-optimal solution for the stochastic ECR problem. Compared with the algorithms proposed in Section 5.2, the main advantage of the the PHA-based algorithm with sequential sampling is that we do not have to determine the sample size in advance.

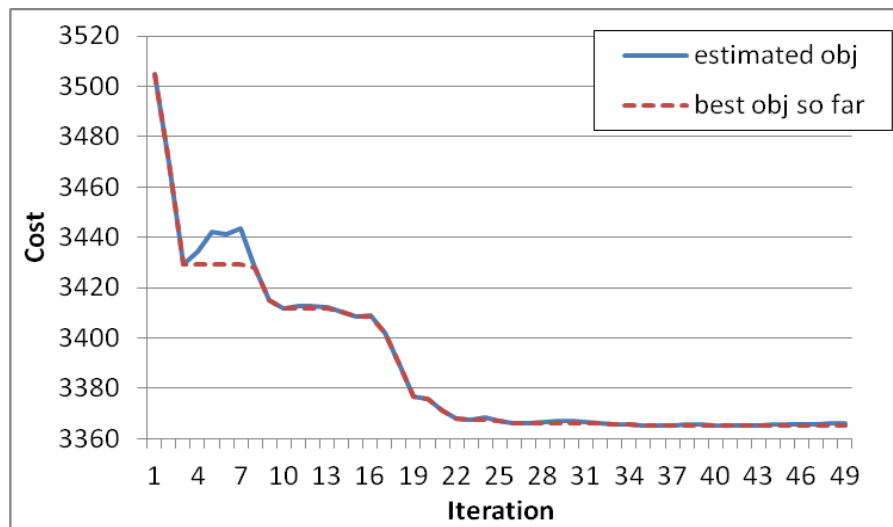


Figure 5.2 The estimated actual cost and the best objective value (case 1)

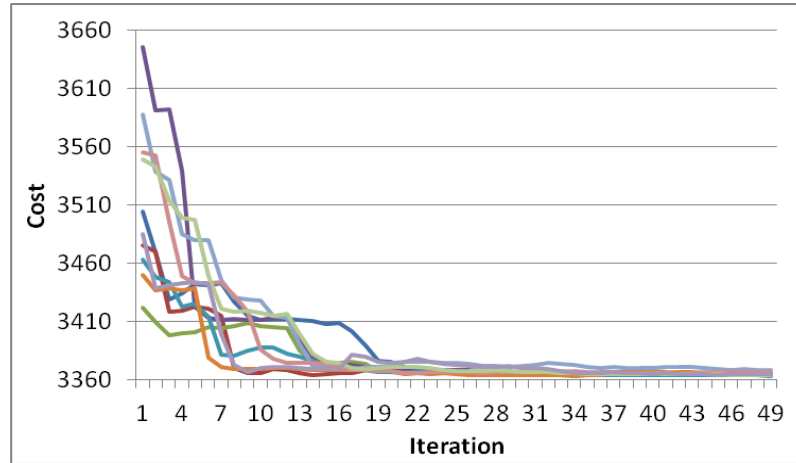


Figure 5.3 The convergence of the PHA-based algorithm with sequential sampling (case 1, 10 replications)

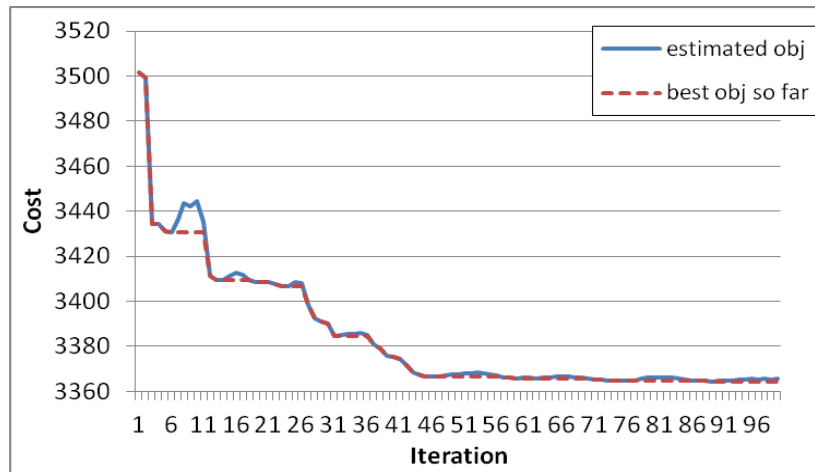


Figure 5.4 The estimated actual cost and the best objective value (case 2)

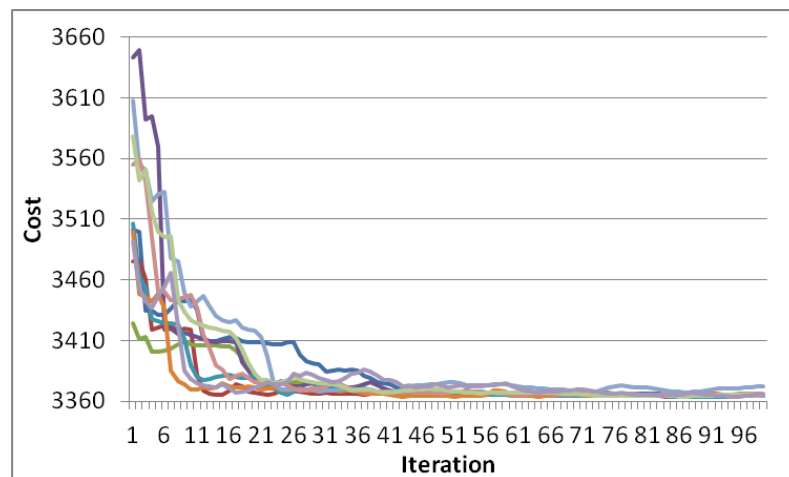


Figure 5.5 The convergence of the PHA-based algorithm with sequential sampling (case 2, 10 replications)

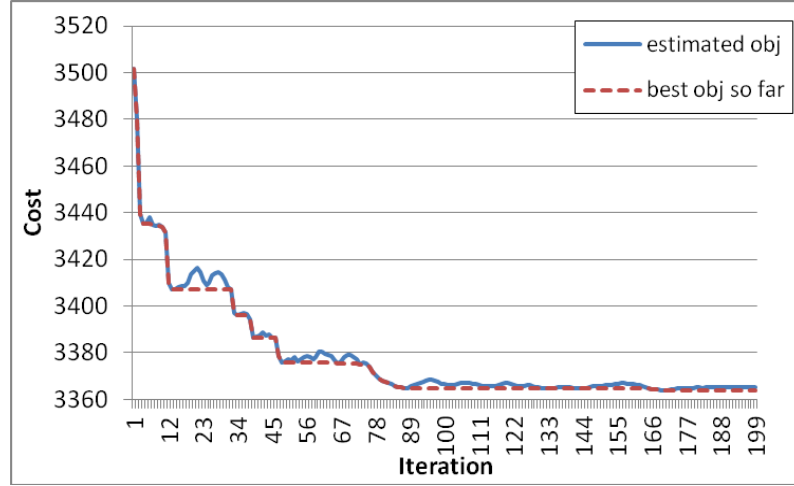


Figure 5.6 The estimated actual cost and the best objective value (case 3)

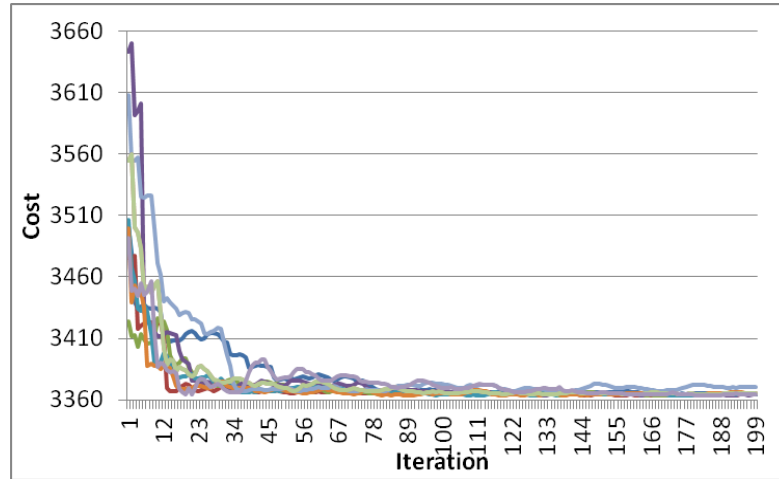


Figure 5.7 The convergence of the PHA-based algorithm with sequential sampling (case 3, 10 replications)

5.4 Summary

To solve the stochastic ECR problem with multiple scenarios, scenario decomposition is considered in this section. The approximate solution of the individual scenario problem is combined to produce an overall solution. Based on the PHA strategy, the feasible solution is calculated iteratively and the best upper bound is kept. In this case, we always have a solution that is best until now at hand.

By generating improving solution, the stochastic ECR problem is optimized.

In this section, we develop two PHA-based algorithms to solve the SAA problem of stochastic ECR. To estimate the performance of these two PHA-based algorithms, we first solve a small-scale case. The result validates that these two PHA-based algorithms can be applied to solve the SAA problem and can provide good approximation of the SAA problem. We also find that the algorithm 2 could provide better approximation. Second, we apply the progressive hedging based algorithms to solve the SAA problem of the large-scale case, which is difficult to solve directly. The results show that these two algorithms can be applied to solve this large-scale network and provide good solutions.

As identifying the appropriate sample size is a crucial problem in the application of the SAA method, we propose a PHA-based algorithm based on sequential sampling. New scenarios are generated and added to the current sample iteratively and thus the sample size does not have to be determined in advance. Computational studies show that PHA-based algorithm based on sequential sampling converges to the optimal solution and could be applied to solve the stochastic ECR problem successfully.

Chapter 6 NON-I.I.D. SAMPLING TO ENHANCE THE SAMPLE AVERAGE APPROXIMATION METHOD

The stochastic Empty Container Repositioning (ECR) problem can be solved by the Sample Average Approximation (SAA) method, in which a sample of random scenarios are generated and then deterministic optimization techniques can be applied to solve the problem based on this sample. In this chapter, we try to enhance the performance of the SAA method by generating more representative samples. In Section 6.1, an introduction on applying non-i.i.d. sampling to enhance the SAA method is presented. In Section 6.2, several sampling schemes are introduced. And then we propose a design which takes the advantages of both Latin hypercube design and supersaturated design. Detailed computational studies on the impact of the sampling methods are presented in Section 6.3.

6.1 Introduction

The SAA problem with multiple scenarios is usually difficult to solve due to its large scale. To solve the large-scale SAA problem efficiently, there are quite a few studies focused on developing algorithms based on scenario decomposition. Another related direction is to generate more representative samples in order to enhance the performance of the SAA method. Independent and identical distribution (i.i.d.) sampling is well studied for construction approximations. The quasi-Monte Carlo method has become popular after the work of Niederreiter (1992) as this method

does not require i.i.d. sampling while works remarkable well. Under this situation, SAA under non-i.i.d. sampling received increasing attention in recent years. The convergence of SAA estimators under general sampling (including i.i.d. and non-i.i.d.), according to our knowledge, was first investigated in Dai et al. (2000). Homem-De-Mello (2008) established the exponential convergence rate under the non-i.i.d. sampling for stochastic optimization problems. Some variance reduction techniques under some appropriate assumptions were theoretically studied in their study, and their results were applied to two sampling schemes, i.e., Latin hypercube sampling and randomized quasi-Monte Carlo. More recently, SAA with general sampling was studied by Xu (2010). The uniform exponential convergence of the sample average of a class of lower semi-continuous random function was first derived in Xu (2010)'s study. Empirically, Freimer et al. (2010) investigated two sampling methods, i.e., Latin hypercube sampling and antithetic variates. The affect of these two sampling schemes on the bias and variance was computationally studied. Results of this study showed that Latin hypercube and antithetic variates methods outperform those under i.i.d. sampling, with Latin hypercube outperforming antithetic variates.

In computer experiments, it is well known that Latin hypercube design achieves maximum stratification in one-dimensional projections. In order to possesses satisfaction in one- and higher dimensional projections, Tang (1993) used orthogonal arrays of strength two or higher to construct Latin hypercubes, i.e., U design. Tang and Qian (2010) used U design to further enhance the accuracy of the SAA and their

theoretical results showed that the SAA with U designs can significantly outperform those with Latin hypercube designs. As U design is based on orthogonal arrays, the main problem with U design is that the required number of experiments increases exponentially with the number of parameters. In this case, it may not be applicable if the system is hectovariate or large.

Based on the established literature, it has been shown that non-i.i.d. sampling could help to enhance the performance of the SAA method. Therefore, we consider applying the SAA method with non-i.i.d. sampling to solve the stochastic ECR problem in this study. On the other hand, although U design has been proved to be able to further enhance the accuracy of SAA, a large number of experiments are required when applying it to our ECR problem which considers a large number of random variables. Inspired by U design, which uses orthogonal arrays to construct Latin hypercubes, we consider constructing Latin hypercubes by using supersaturated design, i.e., designs where the number of experiments is less than the number of variables. As a special branch of design of experiment, the construction of supersaturated design has been well studied. Based on the supersaturated design, we may do only a few experiments (even less than the number of degrees of freedom of the system when that is possible) and still get a satisfying approximation. According to our knowledge, there have been no studies focused on applying supersaturated design to enhance SAA if the scale of the experiment is large and difficult to solve.

The main aim of this study is to enhance the performance of the SAA method for the stochastic ECR problem by generating more representative samples. We

propose a new sampling method, which combines Latin hypercube design and supersaturated design, to enhance the performance of the SAA method for stochastic problem with a large number of random parameters. By using the supersaturated design, we may get better solutions with a few experiments. Specifically, we present a detailed computational study on the performance of the SAA method under several non-i.i.d. sampling schemes, and then compare it with the performance of the standard SAA with i.i.d. sampling. The computational results show that the proposed sampling method is promising.

6.2 Sampling methodology

As explained in Chapter 4, we develop a two-stage stochastic model in rolling horizon policy to deal with the dynamically changing forecasting. ECR decisions are made at the beginning of stage 1 and will be made again when new information is collected. The general formulation of the two-stage stochastic model for ECR is formulated as follows,

$$\min_x c^T x + E_p[Q(x, \xi(\omega))], \text{ subject to } Ax = b, x \geq 0, \quad (6.1)$$

where $Q(x, \xi(\omega))$ is the optimal objective function value of the second-stage problem,

$$Q(x, \xi(\omega)) = \min_y q(\omega)^T y, \text{ subject to } W(\omega)y = h(\omega) + L(\omega)x, y \geq 0, \quad (6.2)$$

$\omega \in \Omega$ denotes a scenario that is unknown when decisions at stage 1 are made, but that is known when the decisions at stage 2 are made, where Ω is the set of all

scenarios. x denotes the first-stage ECR decision, and y denotes the second-stage decision given x and parameters $\xi(\omega)=(q(\omega),W(\omega),h(\omega),L(\omega))$. When the distribution of random parameters is known, the stochastic linear ECR programs can be solved approximately by the SAA method.

$$\hat{z}_N = \min_{x \in X} \{c^T x + \frac{1}{N} \sum_{i=1}^N [Q(x, \xi(\omega^i))]\}, \quad X = \{x: Ax = b, x \geq 0\} \quad (6.3)$$

Sampling should be well-planned when applying the SAA method. The trade-off between the quality of an optimal solution of the SAA problem and the computational effort should be considered. In the standard SAA, \hat{z}_N is estimated with i.i.d. sampling. However, literature shows that non-i.i.d. samplings may help reduce bias as well as variance. In this study, several non-i.i.d. sampling schemes are considered to enhance the performance of the SAA method for the stochastic ECR problem. It has been shown in literature that the performance of Latin hypercube sampling is very well in enhancing SAA. An introduction on Latin hypercube sampling is presented in Section 6.2.1. To generate samples with a small number of scenarios (the number of scenarios is even less than the random variables of the stochastic ECR problem) and still get acceptable solutions, supersaturated design is considered in this study, which is discussed in Section 6.2.2. In Section 6.2.3, we propose a design which combines Latin hypercube design and supersaturated design.

6.2.1 Latin hypercube sampling

It is well known that Latin hypercube design achieves maximum stratification in one-dimensional projections. Sampling with respect to Latin hypercube design is

referred to as Latin hypercube sampling. The idea is to partition the sample space, and the number of sample points on each region should be proportional to the probability of that region. This way we ensure that the number of sampled points on each region will be approximately equal to the expected number of points to fall in that region (McKay et al., 1979). Suppose the random vector of a system consists of m independent components. And we want to draw N scenarios from the random vector. The Latin hypercube sampling operates as follows,

(1). Generate a Latin hypercube $A(N \times m) = (a_{ij})$, where each column of A is a independent random permutation of $1, \dots, N$, with all $N!$ permutations equally probable.

(2). Generate

$$Y^1 \sim U(0, \frac{1}{N}), Y^2 \sim U(\frac{1}{N}, \frac{2}{N}), \dots, Y^N \sim U(\frac{N-1}{N}, 1) \quad (6.4)$$

Based on A , a Latin hypercube design $B = (b_{ij})$ with N runs and m independent random components could be generated by

$$b_{ij} = Y^{a_{ij}} \quad (6.5)$$

(3). For arbitrary distributions, we can apply the inverse-transform technique to generate the desired samplings based on B , which is denoted by $C = (c_{ij})$.

Estimators generated based on the Latin hypercube sampling method are unbiased as it has been shown by McKay et al. (1979) that each random variable c_{ij} in column j (independent random component) has the same distribution.

6.2.2 Supersaturated design

Some practical systems such as the operational ECR problem involve a large number of random variables, e.g., demand and supply at each port, ship weight and space capacity at each voyage. To deal with this problem, we consider applying the supersaturated design to enhance the performance of the SAA method in this study.

A supersaturated design is a design whose run size is not enough for estimating all the main effects, where the number of experiments is less than the number of the degrees of the freedom of the system. In this study, we take the supersaturated design proposed by Ahlinder and Gustafsson (Goury, 2010), which is named as AG design, as an example. It is because of that the AG design owns several good properties and is easy to construct. The procedure to construct two-level AG design with N rows (N experiments) is introduced as follows,

(1). Generate full factorial design of order N , i.e., a design $D(2^N \times N) = (d_{ij})$, where $d_{ij} = 1$ for high level and $d_{ij} = -1$ for low level.

(2). Transpose the design D , and we get a design $E(N \times 2^N) = D^T$.

(3). Remove the first half of the columns and only keep the second half of the columns of E , and we get a design $F(N \times 2^{N-1})$.

(4). Keep the columns that contain a ratio of 1's and -1's equal to 1 (or almost 1 if the columns are of odd length). For the design F , we can keep the columns with $N/2$ 1's and $N/2$ -1's (or $(N-1)/2$ 1's and $(N+1)/2$ -1's when N is odd). And we can finally get the AG design G with N rows and $\left(\frac{N}{N/2}\right)/2$ columns (or

$\binom{N}{(N-1)/2} / 2$ columns when N is odd). For 10 rows, the design has 126 columns.

For a fixed number of variables m , one can get its design by choosing the number of rows that gives a number of columns greater than m . And then delete as many columns as needed. As suggested by Goury (2010), to expect the design to be giving more information on average if you delete a set of columns which is spread out the design matrix. After delete these columns, the desired design $H(N \times m)$ is obtained.

One good property of the AG design is that the saturation increases rather fast with N . By taking the advantage of this property, we are able to generate AG design with a small number of experiments for systems which consider a large number of random components. Another good property is that each column contains a ratio of 1's and -1's equal to 1. It implies that every variable can be fairly evaluated from its lowest level to its highest level. This property is necessary for a stable regression analysis.

6.2.3 The proposed sampling method - Constructing Latin hypercube design by using supersaturated design

Although Latin hypercube design achieves maximum stratification in one-dimensional projections, stratification in higher dimensional projections is not considered as the column for each component is generated independently in Latin hypercube design. To achieve stratification in higher dimensional projections, U

design, which uses orthogonal arrays of strength two or higher to construct Latin hypercubes, was proposed by Tang (1993). And it has been shown that the SAA with U designs can significantly outperform those with Latin hypercube designs (Tang and Qian, 2010). However, the main problem with U design is that it requires a number of experiments increasing exponentially with the number of parameters as the U design is constructed based on orthogonal arrays. Therefore, it may not be applicable if the system is hectovariate or large. Inspired by the idea of U design, we propose to construct Latin hypercube design by using supersaturated design. In this case, by taking the advantages of supersaturated design as well as Latin hypercube design, we may generate designs with a small size of scenarios for systems which consider a large number of variables. Meanwhile, the stratification in one- and higher dimensional projections could also be considered. Suppose we want to construct a design $(N \times m)$ which combines Latin hypercube design and supersaturated design (e.g., the two-level AG design). The procedure of constructing such a design (denoted as AGLH design in this study) is introduced as follows,

- (1). Generate an AG design $H(N \times m)$ according to Section 6.2.2, and randomly permute the rows, columns and symbols of H .
- (2). In each column of H , replace the $N/2$ -1's by a uniform random permutation of $1, \dots, N/2$. And then replace the $N/2$ 1's by a uniform random permutation of $N/2+1, \dots, N$. A design $I(N \times m)$ is obtained.
- (3). Couple $I(N \times m)$ with $U[0, 1]$ random variables according to (6.4) and (6.5). Then apply the inverse-transform technique for arbitrary distributions. The

desired design is denoted as $J(N \times m)$.

As described in Section 6.2.2, to generate the AG design, half of the columns have to be taken away after transposing the full factorial design. In this study, we consider modifying the AG design by removing procedure (3) in Section 6.2.2, while keeping other procedures the same. Thus we can get a modified AG design (denoted as M-AG design in this study) G' with N rows and $\binom{N}{N/2}$ columns (or $\binom{N}{(N-1)/2}$ columns when N is odd). For a fixed number of variables m , we can randomly select m columns from G' , and obtain the desired design $H'(N \times m)$. The main reason to make this modification is to generate samples covering more possible regions. In this case, the bias of the expected perceived cost and the lower bound of the SAA method may be reduced.

The design $J'(N \times m)$ (denoted as M-AGLH design in this study) which combines Latin hypercube design and the modified AG design could be generated according to the similar procedures (1)-(3) of constructing AGLH design.

6.3 Computational studies

This experiment work attempts to examine the performance of SAA method with different sampling schemes for the two-stage stochastic ECR problem. A maritime container transportation network with five ports, three services, and one type of container (twenty-foot standard container) is generated. The transportation network and parameters generation have been introduced in Section 4.4.1 and Appendix B.

There are totally 56 random variables in consideration in this model.

In this section, we empirically compare six sampling schemes, i.e., independent sampling (IS), Latin hypercube sampling (LH), AG sampling (AG), the modified AG sampling (M-AG), a sampling which combines Latin hypercube design and AG design (AGLH), and a sampling which combines Latin hypercube design and the modified AG design (M-AGLH). The sample size N is 10, as 10 is the smallest number of rows that gives the number of columns (with the same number of 1's and -1's) greater than 56 in AG design. The replication number of each sampling method M is 1000. The number of scenarios to evaluate solutions of the SAA problem N' is 1000.

Figure 6.1 is the probability plot of the actual cost estimates, $\hat{z}_{N'}(\hat{x}_N^j), j=1, 2, \dots, M$. For a specific x , y is the probability of the actual cost which is less than x . Figure 6.2 is the box plot of the actual cost estimates. Six methods are compared. The vertical axis is the actual cost estimates of the 1000 replications for each method. The top and bottom of each box are the 25th and 75th percentiles of the samples, respectively. The line in the middle of each box is the sample median and the observations beyond the box are marked as outliers. Based on these figures, we find that all these five non-i.i.d. sampling schemes cause smaller spread of the actual cost estimates compared with the i.i.d. sampling. This result provides evidence that these five non-i.i.d. sampling schemes could help to improve the quality of the solutions of the SAA problems. Especially, AGLH sampling and M-AGLH sampling have the smallest spread of the actual cost estimates. It indicates that robust solutions could be

provided by applying the proposed sampling methods. It may be attributed to that AGLH/M-AGLH takes the advantages of Latin hypercube design as well as the AG/M-AG design. Compared with Latin hypercube sampling, stratification in one- and higher dimensional projections is considered.

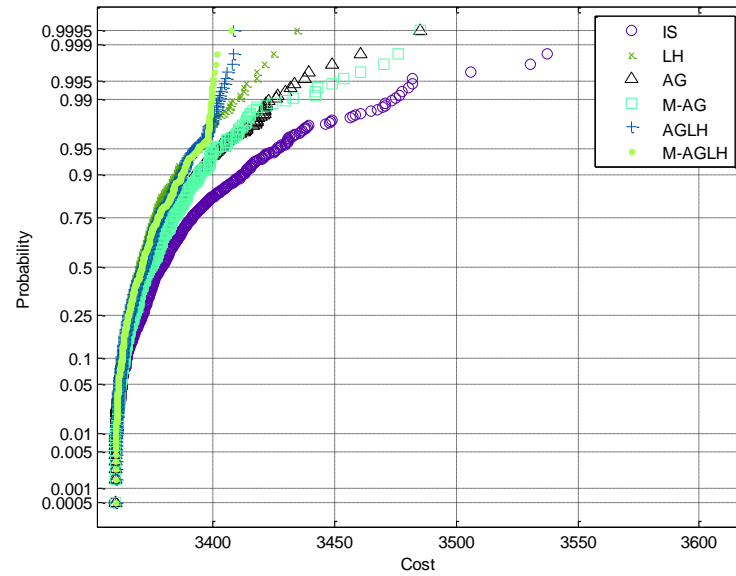


Figure 6.1 Probability plot of the actual cost estimates

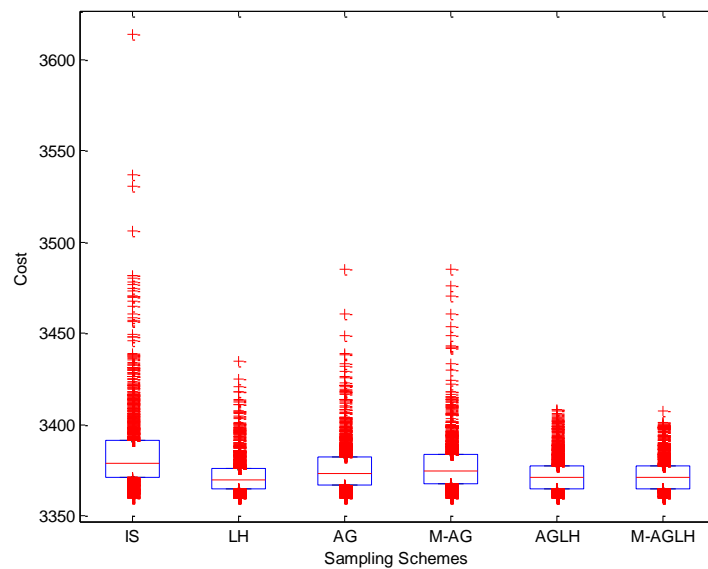


Figure 6.2 Box plot of the actual cost estimates

As explained in Section 4.3, the expected optimal objective value of the SAA problem (4.9) has negative bias, i.e., $E(\hat{z}_N) \leq z^*$, where $E(\hat{z}_N)$ is the expected perceived cost. In this study, the bias reduction of the perceived cost under different sampling schemes is also studied. As the optimal objective value to the true problem (4.1), z^* , is unknown, the bias is unable to calculate. In this study, we use the estimated actual cost of the best solution, which is obtained among all candidate solutions of these six sampling schemes by (4.13), to estimate the true optimal objective value, i.e., 3359.85. Bias is then estimated as the difference between the estimate of $E(\hat{z}_N)$ for a given sampling method and the estimated true optimal objective value. We compare the biases of the non-i.i.d. sampling schemes to the i.i.d. sampling. t -test for bias reduction is conducted in this study. For the t -test (on-sided, $\alpha=0.05$), the null hypothesis is that the expected biases obtained from the non-i.i.d. sampling methods are equal to the expected biases obtained from i.i.d. sampling. Based on our experiments, the bias of the expected perceived cost could be significantly reduced 33.97% by the M-AGLH sampling compared with i.i.d. sampling, while the bias reduction of other non-i.i.d. sampling schemes is statistically insignificant. The good performance of M-AGLH in reducing the bias of the expected perceived cost may be attribute to that the M-AGLH sampling takes the advantages of Latin hypercube sampling as well as the M-AG sampling. Compared to the Latin hypercube design, M-AGLH design considers stratification in one- and higher dimensional projections. However, although AGLH design also considers stratification in one- and higher dimensional projections, the bias of the expected

perceived cost is not significantly improved. One possible reason is that all columns with the same number of 1's and -1's are kept while constructing the M-AG, which is different from the process of removing half of columns while constructing the AG design. Another possible explanation is that we randomly select columns from M-AG design in different replications, while AG samplings for different replications are based on the same AG design (although we conduct row, column and symbol permutation in every replication). In this case, it is possible to cover wider regions and thus reduce the bias of the expected perceived cost.

6.4 Summary

This chapter examines the performance of the SAA method with non-i.i.d. sampling for the stochastic ECR problem. Considering the scale of the SAA problem and the large number of random parameters in the ECR problem, we apply the supersaturated design and aim to get good solutions based on samples with a few scenarios. Computational experiments are designed to examine the impact of different sampling methods for the stochastic ECR problem. Comparing to the i.i.d. sampling, it is found that all the non-i.i.d. samplings in consideration could help to provide robust solutions, thus indicating that the performance of the SAA method could be effectively enhanced. Inspired by the method to constructing U design, a new design is proposed by combining Latin hypercube design and supersaturated design in this study. By considering stratification in one- and higher dimensional

projections, the proposed design has the smallest probability to provide bad solutions. Moreover, the bias of the expected perceived cost could be significantly reduced. By taking the advantage of the supersaturated design, we are able to conduct only a few experiments (even less than the number of degrees of freedom of the system when that is possible) and still get a satisfying approximation. This finding is significant as the proposed sampling method could be applied to systems with a large number of variables.

Although this study has demonstrated the efficiency of several non-i.i.d. samplings for the stochastic ECR problem, the performance of the proposed samplings for other stochastic problems has not been evaluated. To address this problem, future studies are needed to examine the performance of the proposed sampling methods in other stochastic problems. Another possible avenue of future work is to explore new method to construct efficient design in order to enhance the performance of the SAA problem.

Chapter 7 CONCLUSIONS AND FUTURE RESEARCH

This thesis studied operational Empty Container Repositioning (ECR) problem for ocean liners. It contributes to the deterministic and stochastic model formulation and some methodological issues to solve the stochastic ECR problem. In this chapter, we summarize and discuss the main results of our research work as described in previous chapters. Limitations and possible research will also be presented.

7.1 Summary of results

Chapter 3 studied the deterministic ECR problem, aiming to effectively operate empty containers in order to meet demands and to reduce inefficiency. A time space network model which considered the actual operations and constraints of the problems faced by the liner operator was proposed. A real-scale case with 49 ports and 44 actual services was examined in the numerical study. It was found that the handling cost has the largest contribution and is also the most sensitive parameter to the total cost. This result indicates that the container operators should pay more attention to improve the efficiency of empty container handling at ports in order to reduce the ECR cost. We also compared our proposed model with a simple rule which attempted to mimic the actual operation of a shipping liner. The numerical results showed that the proposed model is promising. This proposed model is highly significant because it takes actual services into account and allows flexible

destinations. Thus, it is easy to apply in ECR industry. Moreover, this study has taken the first step in analyzing the empty container transshipment activities at ports. Results on transshipment activities could provide some evidences in indentifying the potential transshipment hubs for empty container transportation.

In Chapter 4, we developed a two-stage stochastic model with uncertain demand, supply, residual ship space capacity, and residual ship weight capacity. The Sample Average Approximation (SAA) method was applied to solve the stochastic ECR problem for the first time. The small optimality gap indicates that the SAA method could provide good solution for the two-stage stochastic ECR problem. Due to the dynamically changing environment and the low forecasting accuracy in container shipping industry, the forecasting has to be adjusted when new information is updated. Thus, the proposed two-stage model in rolling horizon policy is particular promising in dealing with the dynamically changing forecasting. Moreover, by comparing solutions from the deterministic model and the stochastic model, it was found that the operational cost for ECR could be reduced by considering uncertainties. It indicates that the stochastic ECR model which considers uncertain parameters provides more robust decisions, and thus the operation cost for ECR is further reduced.

Methodologies to solve the stochastic ECR problem have been discussed in Chapter 5 and Chapter 6. Chapter 5 considered scenarios decomposition techniques. Several Progressive Hedging Approximation (PHA) -based algorithms were proposed to solve the large-scale SAA problem with multiple scenarios. Our results

showed that the proposed algorithms could successfully be applied to solve the large-scale stochastic ECR problem which is difficult to solve directly. The sequential sampling was also considered in the PHA-based algorithm. Computational studies showed that the PHA-based algorithm based on sequential sampling converges to the optimal solution. The key advantage of the PHA-base algorithm with sequential sampling is that there is no need to determine the sample size in advance. It is noteworthy in that the PHA-based algorithms which were developed to solve our SAA problem for the ECR problem could be easily applied to solve other stochastic programs which consider a large number of scenarios.

In Chapter 6, non-i.i.d. sampling schemes were applied to enhance the performance of the SAA method. By comparing to the i.i.d. sampling, it was found that all the non-i.i.d. samplings in consideration could help to provide robust solutions, thus indicating that the performance of the SAA method could be effectively enhanced. Inspired by the method to constructing U design, a new design was proposed by combining Latin hypercube design and supersaturated design. By considering stratification in one- and higher dimensional projections, it was found that the proposed design has the smallest probability to provide bad solutions. By taking the advantage of the supersaturated design, we are able to conduct only a few experiments (even less than the number of degrees of freedom of the system when that is possible) and still get a satisfying approximation. This work is significant as the proposed sampling method could be applied to systems involving a large number of random variables and each experiment of the system is complex and

time-consuming.

7.2 Possible future research

Despite the contribution described above, the research presented in this thesis has inevitably some limitations. Future research related to the topics reported in this thesis may be carried out in the areas listed below.

In this study, we mainly considered maritime container transportation. Integrated customer demand and supply were considered while detailed inland empty container transposition was not taken into account. It will be interesting to develop operational model which considers both maritime time ECR and inland ECR.

When modeling the problem, we considered ECR for each region separately and did not consider information in other regions when we made ECR decisions in the target region. In the future, this assumption could be relaxed and thus the coordination amongst several regions should be taken into account.

The two-stage stochastic model lives with the assumption that we know the future at the end of the first stage. Therefore, it is unable to handle sequential decisions. Future research should attempt to extend the two-stage problem to multi-stage problem to solve the maritime ECR.

Although we have demonstrated the efficiency of several non-i.i.d. samplings for the stochastic ECR problem, the performance of the proposed samplings for other stochastic problems has not been evaluated. To address this problem, future

studies are needed to examine the performance of the proposed sampling methods in other stochastic problems. Another possible avenue of future work is to explore new method to construct efficient design in order to enhance the performance of the SAA problem.

BIBLIOGRAPHY

- [1] Annual Container Market Review & Forecast 2006/07, Drewry Shipping Consultants LTD: London.
- [2] Acevedo, J., Pistikopoulos, E.N., 1996. Computational studies of stochastic optimization algorithms for process synthesis under uncertainty. *Computers and Chemical Engineering* 20(1), S1-S6.
- [3] Acevedo, J., Pistikopoulos, E.N., 1998. Stochastic optimization based algorithms for process synthesis under uncertainty. *Computers and Chemical Engineering* 22(4-5), 647-671.
- [4] Ahmed, S., Shapiro, A., 2002. The sample average approximation method for stochastic programs with integer recourse. *Optimization Online*, 1-24.
- [5] Bandeira, D.L., Becker, J.L., Borenstein, D., 2009. A DSS for integrated distribution of empty and full containers. *Decision Support Systems* 47(4), 383-397.
- [6] Benders, J.F., 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* 4(3), 238–252.
- [7] Birge, J.R., 1997. Stochastic programming computation and applications. *INFORMS Journal on Computing* 9(2), 111- 133.
- [8] Booth, K.H.V., Cox, D.R., 1962. Some systematic supersaturated designs. *Technometrics* 4(4), 489-495.
- [9] Bulutoglu, D.A., Cheng, C.S., 2004. Construction of E(S2)-optimal

- supersaturated designs. *The Annals of Statistics* 32(4), 1662-1678.
- [10] Butler, N.A., Mead, R., Eskridge, K.M., Gilmour, S.G., 2001. A general method of constructing $E(s_2)$ -optimal supersaturated designs. *Journal of the Royal Statistic Society: Series B, Statistical Methodology* 63(3), 621-632.
- [11] Chang, H., Jula, H., Chassiakos, A., Ioannou, P., 2008. A heuristic solution for the empty container substitution problem. *Transportation Research Part E: Logistics and Transportation Review* 44(2), 203-216.
- [12] Chang, H.S., Fu, M.C., Hu, J., Marcus, S.I., 2007. *Simulation-based algorithms for Markov decision processes* (1st ed.). Berlin: Springer.
- [13] Chen, C., Zeng, Q., 2010. Designing container shipping network under changing demand and freight rates. *Transport* 25(1), 46-57.
- [14] Cheung, R.K., Chen, C.Y., 1998. A two-stage stochastic network model and solution methods for the dynamic empty container allocation problem. *Transportation Science* 32(2), 142-162.
- [15] Cheung, R.K., Powell, W.B., 2000. Shape - a stochastic hybrid approximation procedure for two-stage stochastic programs. *Operations Research* 48(1), 73-79.
- [16] Choong, S.T., Cole, M.H., Kutanoglu, E., 2002. Empty container management for intermodal transportation networks. *Transportation Research Part E: Logistics and Transportation Review* 38(6), 423-438.
- [17] Cordeau, J.F., Soumis, F., Descrosiers, J., 2000. A benders decomposition approach for the locomotive and car assignment problem. *Transportation*

- Science 34(2), 133-149.
- [18] Crainic, T.G., Fu, X., Gendreau, M., Rei, W., Wallace, S.W., 2011. Progressive hedging-based metaheuristics for stochastic network design. *Networks* 58(2), 114-124.
- [19] Crainic, T.G., Gendreau, M., Dejax, P., 1993. Dynamic and stochastic models for the allocation of empty containers. *Operations Research* 41(1), 102-126.
- [20] Dai, L., Chen, C.H., Birge, J.R., 2000. Convergence properties of two-stage stochastic programming. *Journal of Optimization Theory and Applications* 106(3), 489-509.
- [21] Dembo, A., Zeitouni, O., 1998. *Large Deviations Techniques and Applications*. New York: Springer.
- [22] Dong, J.X., Song, D.P., 2009. Container fleet sizing and empty repositioning in liner shipping systems. *Transportation Research Part E: Logistics and Transportation Review* 45(6), 860-877.
- [23] Dong, J.X., Song, D.P., 2012. Quantifying the impact of inland transport times on container fleet sizing in liner shipping services with uncertainties. *OR Spectrum* 34(1), 155-180.
- [24] Du, Y., Hall, R., 1997. Fleet sizing and empty equipment redistribution for center-terminal transportation networks. *Management Science* 43(2), 145-157.
- [25] Erera, A.L., Morales, J.C., Savelsbergh, M., 2005. Global intermodal tank container management for the chemical industry. *Transportation Research Part E: Logistics and Transportation Review* 41(6), 551-566.

- [26] Erera, A.L., Morales, J.C., Savelsbergh, M. 2009. Robust optimization for empty repositioning problems. *Operations Research* 57 (2), 468-483.
- [27] Ermoliev, Y., 1988. Stochastic quasigradient methods. In Ermoliev, Y., Wets. Y. (Eds), *Numerical methods in stochastic programming* (141-185). Berlin: Springer-Verlag.
- [28] Feng, C.M., Chang, C.H., 2008. Empty container reposition planning for intra-Asia liner shipping. *Maritime Policy and Management* 35(5), 469-489.
- [29] Francesco, M.D., Crainic, T.G., Zuddas, P., 2009. The effect of multi-scenario policies on empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review* 45(5), 758-770.
- [30] Freimer, M.B., Linderoth, J.T., Thomas, D., 2010. The impact of sampling methods on bias and variance in stochastic linear programs. *Computational Optimization and Applications* 51(1), 51-75.
- [31] Fu, M.C., 2002. Optimization for simulation: theory vs. practice, *INFORMS Journal on Computing* 14(3), 192-215.
- [32] Fu, M.C., Glover, F.W., April, J., 2005. Simulation optimization: a review, new developments, and applications. *Proceedings of the 2005 Winter Simulation Conference*, 83-95.
- [33] Geyer, C.J., Thompson, E.A., 1992. Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society Series B (Methodological)* 54(3), 657-699.
- [34] Gorman, M.F., 2002. Pricing and market mix optimization in freight

- transportation. *Journal of the Transportation Research Forum* 56 (1), 135-148.
- [35] Goury, O., 2010. One efficient black box optimization of systems defined by 100 or more parameters. Thesis for the degree of Master of Science. Chalmers University of Technology.
- [36] Haugen, K.K., Løkketangen, A., Woodruff, D.L., 2001. Progressive hedging as a meta-heuristic applied to stochastic lot-sizing. *European Journal of Operational Research* 132(1), 116-122.
- [37] Helgason, T., Wallace, S.W., 1991. Approximate scenario solutions in the progressive hedging algorithm. *Annals of Operations Research* 31(1), 425-444.
- [38] Higle, J.L., 1998. Variance reduction and objective function evaluation in stochastic linear programs. *INFORMS Journal on Computing* 10(2), 236-247.
- [39] Higle, J.L., Sen, S., 1991. Stochastic decomposition: an algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research* 16(3), 650–669.
- [40] Homem-De-Mello, T., 2008. On rates of convergence for stochastic optimization problems under non-independent and identically distributed sampling. *SIAM Journal on Optimization* 19(2), 524-551.
- [41] Hurtado, J.E., Barbat, A.H., 1996. Improved stochastic linearization method using mixed distributions. *Structural Safety* 18(1), 49-62.
- [42] Hvattum, L.M., Løkketangen, A., 2009. Using scenario tree and progressive hedging for stochastic inventory routing problems. *Journal of Heuristics* 15(6),

527-557.

- [43] Imai, A., Shintani, K., Papadimitriou, S., 2009. Multi-port vs. hub-and-spoke port calls by container ships. *Transportation Research Part E: Logistics and Transportation Review* 45(5), 740-757.
- [44] Infanger, G., 1994. *Planning under uncertainty: solving large scale stochastic linear programs*. Denver, MA: Boyd and Fraser.
- [45] Jula, H., Chassiakos, A., Ioannou, P., 2006. Port dynamic empty container reuse. *Transportation Research Part E: Logistics and Transportation Review* 42(1), 43-60.
- [46] Kim, K.J., Diwekar, U.M. 2002. Efficient combinatorial optimization under uncertainty. 1. Algorithmic development. *Industrial & Engineering Chemistry Research* 41(5), 1276-1284.
- [47] Kleywegt, A.J., Shapiro, A., Homem-De-Mello, T., 2002. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12(2), 479-502.
- [48] Korf, L.A., 1998. *Approximation and solution schemes for stochastic dynamic optimization problems*. Doctoral Dissertation, University of California Davis.
- [49] Lai, K.K., Lam, K., Chan, W.K., 1995. Shipping container logistics and allocation. *The Journal of the Operational Research Society* 46(6), 687-697.
- [50] Lam, S.W., Lee, L.H., Tang, L.C., 2007. An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C: Emerging Technologies* 15(4), 265-277.

- [51] Lasdon, L.S., 2002. Optimization theory for large systems. Mineola, New York: Dover Publications, Inc..
- [52] Li, J.A., Leung, S.C.H., Wu, Y., Liu, K., 2007. Allocation of empty containers between multi-ports. *European Journal of Operational Research* 182(1), 400-412.
- [53] Li, J.A., Liu, K., Leung, S.C.H., Lai, K.K., 2004. Empty container management in a port with long-run average criterion. *Mathematical and Computer Modelling* 40(1-2), 85-100.
- [54] Lin, D.K.J., 1993. A new class of supersaturated designs. *Technometrics* 35(1), 28-31.
- [55] Lin, D.K.J., 1995. Generating systematic supersaturated designs. *Technometrics* 37(2), 213–225.
- [56] Linderoth, J., Shapiro, A., Wright, S., 2006. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research* 142(1), 215-241.
- [57] Løkketangen, A., Woodruff, D.L., 1996. Progressive hedging and tabu search applied to mixed integer (0, 1) multistage stochastic programming. *Journal of Heuristics* 2(2), 111-128.
- [58] Mak, W.K., Morton, D.P., Wood, R.K., 1999. Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters* 24(1-2), 47-56.
- [59] Marar, A., 2002. Information representation in large-scale resource allocation

- problems: theory, algorithms and applications. Ph.D. thesis, Princeton University, Princeton, NJ.
- [60] McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* 21(2), 239-245.
- [61] Moon, I.K., Ngoc, A.D.D., Hur, Y.S., 2010. Positioning empty containers among multiple ports with leasing and purchasing considerations. *OR Spectrum* 32(3), 765-786.
- [62] Nguyen, N.K., 1996. An algorithmic approach to constructing supersaturated designs. *Technometrics* 38(1), 69-73.
- [63] Niederreiter, H., 1992. *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia, PA: SIAM.
- [64] Norkin, V.I., Ermoliev, Y.M., Ruszczyński, A., 1998a. On optimal allocation of indivisibles under uncertainty. *Operations Research* 46(3), 381-395.
- [65] Norkin, V.I., Pflug, G.Ch., Ruszczyński, A., 1998b. A branch and bound method for stochastic global optimization. *Mathematical Programming* 83(3), 425-450.
- [66] Novak, Z., Kravanja, Z., 1999. Mixed-integer nonlinear programming problem process synthesis under uncertainty by reduced dimensional stochastic optimization. *Industrial & Engineering Chemistry Research* 38(7), 2680-2698.
- [67] Olivo, A., Zuddas, P., Francesco, M.D., Manca, A., 2005. An operational

- model for empty container management. *Maritime Economics & Logistics* 7(3), 199-222.
- [68] Pennanen, T., 2005. Epi-convergent discretizations of multistage stochastic programs. *Mathematics of Operations Research* 30(1), 245-256.
- [69] Pennanen, T., Kallio, M., 2006. A splitting method for stochastic programs. *Annals of Operations Research* 142(1), 259-268.
- [70] Pintarič, Z.N., Kravanja, Z., 2000. The two-level strategy for MINLP synthesis of process flowsheets under uncertainty. *Computers and Chemical Engineering*, 24(2-7), 195-201.
- [71] Plambeck, E.L., Fu, B.R., Robinson, S.M., Suri, R., 1996. Sample-path optimization of convex stochastic performance functions. *Mathematical Programming* 75(2), 137-176.
- [72] Powell, W.B., 2007. *Approximate dynamic programming: solving the curses of dimensionality*. Hoboken, New Jersey: John Wiley & Sons, Inc..
- [73] Powell, W.B., 2009. What you should know about approximate dynamic programming. *Naval Research Logistics* 56(3), 239-249.
- [74] Powell, W.B., Jaillet, P., Odoni, A., 1995. Stochastic and dynamic networks and routing. In Ball, M.O., Magnanti, T.L., Monma, C.L., Nemhauser, G.L. (Eds.), *Handbook in operations research and management science, Volume 8: Network Routing* (141-295). Amsterdam: Elsevier.
- [75] Powell, W.B., Marar, A., Kulkarni, S., 2006. Capturing expert knowledge in resource allocation problems through low-dimensional patterns. *IIE*

- Transactions 38, 159-172.
- [76] Powell, W.B., Topaloglu, H., 2005. Fleet Management. In Wallace, S.W., Ziemba, W.T. (Eds.), Applications of Stochastic Programming (185-216), Volume 5 MPS-SIAM series on optimization. Philadelphia: The society for industrial and applied mathematics and the mathematical programming society.
- [77] Rockafellar, R.T., Wets, R.J., 1991. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research* 16(1), 119-147.
- [78] Rodrigue, J.P., Comtois, C., Slack, B., 2009. International trade and freight distribution. *The geography of transport systems* (2nd Edition), Chapter 5. New York: Routledge.
- [79] Rubinstein, R.Y., Shapiro, A., 1990. Optimization of static simulation models by the score function method. *Mathematics and Computers in Simulation* 32(4), 373-392.
- [80] Ruszczyński, A., 1997. Decomposition methods in stochastic programming. *Mathematical Programming* 79 (1-3), 333-353.
- [81] Ruszczyński, A., Shapiro, A. (Eds.), 2003. Stochastic Programming. Vol. 10 of *Handbooks in Operations Research and Management Science*. Amsterdam: Elsevier.
- [82] Salinger, D.H., 1997. A splitting algorithm for multistage stochastic programming with application to hydropower scheduling. Doctoral

- Dissertation, Dept. of Applied Math., University of Washington.
- [83] Sanger, T.D., 1991. A tree-structured adaptive network for function approximation in high-dimensional spaces. *Neural Networks, IEEE Transactions* 2(2), 285-293.
- [84] Santoso, T., Ahmed, S., Goetschalckx, M., Shapiro, A., 2005. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research* 167(1), 96-115.
- [85] Shapiro, A., 1996. Simulation-based optimization: convergence analysis and statistical inference. *Stochastic Models* 12(3), 425-454.
- [86] Shapiro, A., Homem-de-Mello, T., 1998. A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming* 81(3), 301–325.
- [87] Shen, W.S., Khoong, C.M., 1995. A DSS for empty container distribution planning. *Decision Support Systems* 15(1), 75-82.
- [88] Shintani, K., Imai, A., Nishimura, E., Papadimitriou, S., 2007. The container shipping network design problem with empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review* 43(1), 39-59.
- [89] Silva, A.D., Elias, M., Andraud, C., Lafait, J., 2003. Comparison of the auxiliary function method and the discrete-ordinate method for solving the radiative transfer equation for light scattering. *Optical Society of America Journal* 20(12), 2321-2329.

- [90] Simão, H.P., Day, J., George, A.P., Gifford, T., Nienow, J., Powell, W.B., 2009. An approximate dynamic programming algorithm for large-scale fleet management: a case application. *Transportation Science* 43(2), 178-197.
- [91] Soize, C., 1995. Stochastic linearization method with random parameters for SDOF nonlinear dynamical systems: prediction and identification procedures. *Probabilistic Engineering Mechanics* 10(3), 143-152.
- [92] Song, D.P., 2005. Optimal threshold control of empty vehicle redistribution in two depot service systems. *Automatic Control, IEEE Transactions on* 50(1), 87-90.
- [93] Song, D.P., 2007. Characterizing optimal empty container reposition policy in periodic-review shuttle service systems. *Journal of the Operational Research Society* 58, 122-133.
- [94] Song, D.P., Carter, J., 2008. Optimal empty vehicle redistribution for hub-and-spoke transportation systems. *Naval Research Logistics* 55(2), 156-171.
- [95] Song, D.P., Carter, J., 2009. Empty container repositioning in liner shipping. *Maritime Policy and Management* 36(4), 291-307.
- [96] Song, D.P., Dong, J.X., 2008. Empty container management in cyclic shipping routes. *Maritime Economics & Logistics* 10(4), 335-361.
- [97] Song, D.P., Dong, J.X., 2011. Effectiveness of an empty container repositioning policy with flexible destination ports. *Transport Policy* 18(1), 92-101.

- [98] Song, D.P., Earl, C.F., 2008. Optimal empty vehicle repositioning and fleet-sizing for two-depot service systems. *European Journal of Operational Research* 185(2), 760-777.
- [99] Spall, J.C., 2003. Introduction to stochastic search and optimization: estimation, simulation, and control. Hoboken, New Jersey: John Wiley & Sons, Inc..
- [100] Steenken, D., Voß, S., Stahlbock, R., 2004. Container terminal operation and operations research - a classification and literature review. *OR Spectrum* 26(1), 3-49.
- [101] Tang, B., 1993. Orthogonal array-based Latin hypercubes. *Journal of the American Statistical Association* 88(424), 1392-1397.
- [102] Tang, Q., Qian, P.Z.G., 2010. Enhancing the sample average approximation method with U design. *Biometrika*, 1-14.
- [103] Topaloglu, H., Powell, W.B., 2006. Dynamic programming approximations for stochastic time-staged integer multicommodity-flow problems. *INFORMS Journal on Computing* 18(1), 31-42.
- [104] Topaloglu, H., Powell, W.B., 2007. Incorporating pricing decisions into the stochastic dynamic fleet management problem. *Transportation Science* 41 (3), 281-301.
- [105] Van Slyke, R.M., Wets R., 1969. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics* 17(4), 638-663.

- [106] Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G., Shapiro, A., 2003. The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications* 24(2-3), 289-333.
- [107] Wallace, S.W., Ziemba, W.T., 2005. Applications of stochastic programming. Volume 5 MPS-SIAM series on optimization. Philadelphia: The society for industrial and applied mathematics and the mathematical programming society.
- [108] Wei, J., Realff, M.J., 2004. Sample average approximation methods for stochastic MINLPs. *Computers & Chemical Engineering* 28(3), 333-346.
- [109] White, W.W., 1972. Dynamic transshipment networks: an algorithm and its application to the distribution of empty containers. *Networks* 2(3), 211-236.
- [110] Wu, Z., Bai, F., Mammadov, M., 2008. An auxiliary function method for constrained systems of nonlinear equations. Institute of Mathematics and Informatics, Vilnius Gediminas Technical University.
- [111] Xu, H., 2010. Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming. *Journal of Mathematical Analysis and Applications* 368, 692-710.
- [112] Ye, H.Q., Yuan, X.M., Liu, X., 2007. A tactical planning model for liner shipping companies: managing container flow and ship deployment jointly. School of Business, National University of Singapore.
- [113] Zhang, R., Yun, W.Y., Kopfer, H., 2010. Heuristic-based truck scheduling for

- inland container transportation. *OR Spectrum* 32(3), 787-808.
- [114] Zhang, R., Yun, W.Y., Moon, I., 2009. A reactive tabu search algorithm for the multi-depot container truck transportation problem. *Transportation Research Part E: Logistics and Transportation Review* 45(6), 904-914.
- [115] Zhang, W., Wang, G., Xing, Z., Wittenburg, L., 2005. Distributed stochastic search and distributed breakout: properties, comparison and applications to constraint optimization problems in sensor networks. *Artificial Intelligence* 161(1-2), 55-87.
- [116] Zhou, W.H., Lee, C.Y., 2009. Pricing and competition in a transportation market with empty equipment repositioning. *Transportation Research Part B: Methodological* 43(6), 677-691.

APPENDICES

Appendix A: The explicit form of the two-stage model P2

Besides the concise formulation, the explicit form of the two-stage model **P2** is also presented to help reader understand the two-stage model. Suppose the planning horizon is T , the length of stage 1 is T_1 , where $T_1 < T$.

P2-Stage 1:

$$\begin{aligned}
 \min \quad & \sum_{t=1,2,\dots,T_1} \sum_{k \in K} \sum_{i \in P} c_{t,i,k}^u \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, p_{v,s}=i, t \in A_{v,s}\}} u_{t,s,v,k} \\
 & + \sum_{t=1,2,\dots,T_1} \sum_{i \in P} \sum_{k \in K} (c_{t,i,k}^y y_{t,i,k} + c_{t,i,k}^z z_{t,i,k}) \\
 & + \sum_{t=1,2,\dots,T_1} \sum_{k \in K} \sum_{i \in P} c_{t,i,k}^w \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, p_{v,s}=i, t \in D_{v,s}\}} w_{t,s,v,k} \\
 & + \sum_{t=1,2,\dots,T_1} \sum_{k \in K} \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, t \in D_{v,s}\}} c_{t,s,v,k}^x x_{t,s,v,k} \\
 & + E_p[Q(x_1, \xi(\omega))]
 \end{aligned} \tag{A.1}$$

Subject to

$$\sum_{k \in K} (g_k \times x_{t,s,v,k}) \leq \gamma_{t,s,v} \quad \forall (v,s,t) \in \{(v,s,t) | v \in V, s \in S_v, t \in D_{v,s}, t \leq T_1\} \tag{A.2}$$

$$\sum_{k \in K} (h_k \times x_{t,s,v,k}) \leq \sigma_{t,s,v} \quad \forall (v,s,t) \in \{(v,s,t) | v \in V, s \in S_v, t \in D_{v,s}, t \leq T_1\} \tag{A.3}$$

$$\begin{aligned}
 x_{t-b_{v,s},s,v,k} - u_{t,s+1,v,k} + w_{t+d_{v,s+1},s+1,v,k} &= x_{t+d_{v,s+1},s+1,v,k} \\
 \forall k \in K, \forall (v,s,t) \in \{(v,s,t) | v \in V, s \in S_v, t \in A_{v,s+1}, t \leq T_1\}
 \end{aligned} \tag{A.4}$$

$$x_{t-b_{v,s},s,v,k} \geq u_{t,s+1,v,k} \quad \forall k \in K, \forall (v,s,t) \in \{(v,s,t) | v \in V, s \in S_v, t \in A_{v,s+1}, t \leq T_1\} \tag{A.5}$$

$$\begin{aligned}
 y_{t-1,i,k} + \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, p_{v,s}=i, t \in A_{v,s}\}} u_{t,s,v,k} + z_{t,i,k} + \theta_{t,i,k} - \sum_{(s,v) \in \{(s,v) | v \in V, s \in S_v, p_{v,s}=i, t \in D_{v,s}\}} w_{t,s,v,k} \\
 - \psi_{t,i,k} = y_{t,i,k} \quad \forall k \in K, \forall i \in P, \forall t = 1, 2, \dots, T_1
 \end{aligned} \tag{A.6}$$

$$y_{t,i,k} = v_1 \quad \forall k \in K, \forall i \in P, t = T_1 \tag{A.7}$$

$$x_{t,s,v,k} = v_2 \quad \forall k \in K, \forall (v,s,t) \in \{(v,s,t) \mid v \in V, s \in S_v, t \in D_{v,s}, t \leq T_1, t + b_{v,s} > T_1\} \quad (\text{A.8})$$

$$x_{t-b_{v,s},s,v,k} - u_{t,s+1,v,k} = v_3 \quad (\text{A.9})$$

$$\forall k \in K, \forall (v,s,t) \in \{(v,s,t) \mid v \in V, s \in S_v, t \in A_{v,s+1}, t \leq T_1, t + d_{v,s+1} > T_1\}$$

$$u_{t,s,v,k}, w_{t,s,v,k}, x_{t,s,v,k}, y_{t,i,k}, z_{t,i,k} \geq 0 \quad (\text{A.10})$$

$$\forall k \in K, \forall i \in P, \forall t = 1, 2, \dots, T_1, \forall (v,s) \in \{(v,s) \mid v \in V, s \in S_v\}$$

x_1 : Decisions at stage 1, including $u_{t,s,v,k}$, $w_{t,s,v,k}$, $x_{t,s,v,k}$, $y_{t,i,k}$, and $z_{t,i,k}$ at stage 1;

v_1 : Empty container inventory at a port at the end of stage 1;

v_2 : Empty container inventory on vessels when these vessels are travelling at the end of stage 1;

v_3 : Empty container inventory on vessels when these vessels are staying at a port at the end of stage 1;

v : The vector of ending container states of stage 1. It is the empty container inventory at each port and at each vessel at the end of stage 1. $v = [v_1, v_2, v_3]$.

P2-Stage 2: For a realized scenario ω , we have

$$\begin{aligned} Q[x_1, \xi(\omega)] = & \min \sum_{t=T_1+1, \dots, T} \sum_{k \in K} \sum_{i \in P} c_{t,i,k}^u \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s}=i, t \in A_{v,s}\}} u_{t,s,v,k}(\omega) \\ & + \sum_{t=T_1+1, \dots, T} \sum_{i \in P} \sum_{k \in K} (c_{t,i,k}^y y_{t,i,k}(\omega) + c_{t,i,k}^z z_{t,i,k}(\omega)) \\ & + \sum_{t=T_1+1, \dots, T} \sum_{k \in K} \sum_{i \in P} c_{t,i,k}^w \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s}=i, t \in D_{v,s}\}} w_{t,s,v,k}(\omega) \\ & + \sum_{t=T_1+1, \dots, T} \sum_{k \in K} \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, t \in D_{v,s}\}} c_{t,s,v,k}^x x_{t,s,v,k}(\omega) \end{aligned} \quad (\text{A.11})$$

Subject to

$$\sum_{k \in K} (g_k \times x_{t,s,v,k}(\omega)) \leq \gamma_{t,s,v}(\omega) \quad (\text{A.12})$$

$$\forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in D_{v,s}, T_1 < t \leq T\}$$

$$\sum_{k \in K} (h_k \times x_{t,s,v,k}(\omega)) \leq \sigma_{t,s,v}(\omega) \quad (A.13)$$

$$\forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in D_{v,s}, T_1 < t \leq T\}$$

$$x_{t-b_{v,s},s,v,k}(\omega) - u_{t,s+1,v,k}(\omega) + w_{t+d_{v,s+1},s+1,v,k}(\omega) = x_{t+d_{v,s+1},s+1,v,k}(\omega) \quad (A.14)$$

$$\forall k \in K, \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in A_{v,s+1}, T_1 < t \leq T\}$$

$$x_{t-b_{v,s},s,v,k}(\omega) \geq u_{t,s+1,v,k}(\omega) \quad (A.15)$$

$$\forall k \in K, \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in A_{v,s+1}, T_1 < t \leq T\}$$

$$y_{t-1,i,k}(\omega) + \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s}=i, t \in A_{v,s}\}} u_{t,s,v,k}(\omega) + z_{t,i,k}(\omega) - \sum_{(s,v) \in \{(s,v) \mid v \in V, s \in S_v, p_{v,s}=i, t \in D_{v,s}\}} w_{t,s,v,k}(\omega) - \psi_{t,i,k}(\omega) + \theta_{t,i,k}(\omega) = y_{t,i,k}(\omega) \quad (A.16)$$

$\forall k \in K, \forall i \in P, \forall t = T_1 + 1, \dots, T$

$$y_{t,i,k}(\omega) = v_1(x_1) \quad \forall k \in K, \forall i \in P, t = T_1 \quad (A.17)$$

$$x_{t,s,v,k}(\omega) = v_2(x_1) \quad (A.18)$$

$$\forall k \in K, \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in D_{v,s}, t \leq T_1, T_1 < t + b_{v,s} \leq T\}$$

$$x_{t,s+1,v,k}(\omega) - w_{t,s+1,v,k}(\omega) = v_3(x_1) \quad (A.19)$$

$$\forall k \in K, \forall (v, s, t) \in \{(v, s, t) \mid v \in V, s \in S_v, t \in D_{v,s+1}, t - d_{v,s+1} \leq T_1, T_1 < t \leq T\}$$

$$u_{t,s,v,k}(\omega), w_{t,s,v,k}(\omega), x_{t,s,v,k}(\omega), y_{t,i,k}(\omega), z_{t,i,k}(\omega) \geq 0 \quad (A.20)$$

$$\forall k \in K, \forall i \in P, \forall t = T_1 + 1, \dots, T, \forall (v, s) \in \{(v, s) \mid v \in V, s \in S_v\}$$

$v_1(x_1)$: Initial empty container inventory at a port of stage 2 given x_1 ;

$v_2(x_1)$: Initial empty container inventory on vessels when these vessels are travelling

at the beginning of stage 2 given x_1 ;

$v_3(x_1)$: Initial empty container inventory on vessels when these vessels are staying at

a port at the beginning of stage 2 given x_1 ;

$v(x_1)$: The vector of initial container states of stage 2 given x_1 . $v(x_1) = [v_1(x_1), v_2(x_1)]$,

$v_3(x_1)]$.

The stage 1 model and stage 2 model are connected by the container flow between the two stages. Given x_1 , the initial container states of stage 2, $v(x_1)$, should equal to the ending container states of stage 1, v .

Appendix B: Data generation and cost parameter of the small-scale case

The demand and supply in stage 2 are generated based on normal distribution. Demand of these five ports is generated according to $N(10, 5^2)$, $N(15, 7.5^2)$, $N(10, 5^2)$, $N(15, 7.5^2)$, and $N(20, 10^2)$ respectively, and supply of these five ports is generated according to $N(5, 2.5^2)$, $N(15, 7.5^2)$, $N(5, 2.5^2)$, $N(5, 2.5^2)$, and $N(5, 2.5^2)$ respectively. The total ship space capacity (denoted by TSS) of $S1$, $S2$ and $S3$ is set to be 50 TEU (Twenty-foot Equivalent Unit), 50 TEU and 25 TEU respectively. The total ship weight capacity (denoted by TSW) is set to be 500 DWT (Dead Weight Tonnage), 500 DWT and 250 DWT respectively. To generate the available ship capacity for empty container, we first define a factor γ to represent the percentage of available ship space capacity, where γ follows the normal distribution $N(0.35, 0.2^2)$. So the residual ship space capacity (denoted by RSS) of voyages at service i are generated according to

$$RSS_i = TSS_i \times N(0.35, 0.2^2), \quad RSS_i \in [0, TSS] \quad (\text{B.21})$$

Considering the high positive correlation between the laden ship space capacity

and the laden ship weight capacity, we define a factor ε to represent the correlation between the laden ship space capacity and the laden ship weight capacity, where ε follows the normal distribution $(13, 2^2)$. The residual ship weight capacity (denoted by RSW) of voyages at service i are generated according to

$$RSW_i = TSW_i - N(13, 2^2)[1 - N(0.35, 2^2)]TSS_i, \quad RSW_i \in [0, TSW] \quad (\text{B.22})$$

The relevant cost parameters are shown in Table B.1.

Table B.1 Cost parameters of ECR problem (small-scale case)

Port	1	2	3	4	5
Handling cost(\$/unit)	10	9	8	13	10
Storage cost(\$/unit/day)	0.2	0.2	0.2	0.3	0.2
Penalty cost(\$/unit)	50	45	40	65	50
Transport cost (\$/unit/voyage)			2		